Universidade de Aveiro
2012
Departamento de Eletrónica, Telecomunicações e Informática

Ilídio Fernando de Castro Oliveira

Uma rede telemática para a prestação regional de cuidados de saúde

A telematic platform towards regional connected healthcare

Universidade de Aveiro
2012
Departamento de Eletrónica, Telecomunicações e Informática

Ilídio Fernando de Castro Oliveira

Uma rede telemática para a prestação regional de cuidados de saúde

A telematic platform towards regional connected healthcare

Tese apresentada à Universidade de Aveiro para cumprimento dos requisitos necessários à obtenção do grau de Doutor em Engenharia Informática, realizada sob a orientação científica do Doutor João Paulo T. Silva Cunha, Professor Associado com Agregação da Faculdade de Engenharia da Universidade do Porto, e do Doutor António M. Melo de Sousa Pereira, Professor Catedrático do Departamento de Eletrónica, Telecomunicações e Informática da Universidade de Aveiro.

Este trabalho foi parcialmente apoiado pelo programa “AveiroDigital” da iniciativa “PortugalDigital” do Governo Português, com fundos do POSI.

Este trabalho foi parcialmente desenvolvido no contexto do projeto GERES-med, apoiado financeiramente pela Fundação para a Ciência e Tecnologia (GRID/GRI/81819/2006).

Health informatics is a tool for better health care and research.

This work is dedicated to those who operate the tool.

o júri

presidente

Prof. Doutora Maria Celeste da Silva do Carmo
Professora Catedrática da Universidade de Aveiro

vogais

Prof. Doutor António Manuel Melo de Sousa Pereira (coorientador)
Professor Catedrático da Universidade de Aveiro

Prof. Doutor João Paulo Trigueiros da Silva Cunha (orientador)
Professor Associado com Agregação da Faculdade de Engenharia da Universidade do Porto

Prof. Doutora Inês de Castro Dutra
Professora Auxiliar da Faculdade de Ciências da Universidade do Porto

Prof. Doutor Ricardo Cruz Correia
Professor Auxiliar da Faculdade de Medicina da Universidade do Porto

Prof. Doutor José Carlos Nascimento
Professor Auxiliar da Escola de Engenharia da Universidade do Minho

agradecimentos

acknowledgements

This is the only section in which the first-person singular is use; in the next chapters, I will prefer the use of 'we'. Besides a style option, it is a statement that I was fortunate to collaborate with several people, to whom I am in debt.

I would like to start this work by acknowledging the support from several teams:

- I am grateful to *Instituto de Engenharia Electrónica e Telecomunicações* (IEETA), represented by its Directors, which provided the material conditions for this work to take place. I have found in this 'house of science' many examples of high standards in research, impelling one to always do better.
- I am grateful to all the people involved in the *Rede Telemática da Saúde* project activities. This was a valuable opportunity to work with professionals from very different backgrounds. I extend my acknowledgement to the organizations involved, represented by the 'liaison' staff: César Telmo Costa (HIP), Amândio Pereira (HDA), Alberto Neiva (ex-SRSAv) and Isabel Cruz (UA).
- I am grateful to all friends at IEETA, for the everyday companionship and fruitful exchange of opinions.
- I am grateful to Professor Orlando Belo (UM) for the words of advice and availability.

This work would never have happen without the vision, effort and interest of Professor João Paulo Cunha. I owe him the opportunity to learn from and contribute to his health informatics group, the opportunity to play a technical coordination role in the RTS project, and the opportunity to develop this work under his supervision.
The interest in the integration of biomedical knowledge across different domains came from the vision of Professor A. Sousa Pereira. I am grateful for the important role he played concerning the material conditions and finalization of this work. Working with him has been a truly enriching experience, much beyond the technical level.

Finally, to those I do not mention here by name, I assure you I have your help present by heart (family included).

palavras-chave

Tecnologias de informação em saúde, redes regionais de saúde, Registo de Saúde Eletrónico, integração de sistemas, computação Grid, HealthGrid.

resumo

As tecnologias de informação e comunicação na área da saúde não são só um instrumento para a boa gestão de informação, mas antes um fator estratégico para uma prestação de cuidados mais eficiente e segura. As tecnologias de informação são um pilar para que os sistemas de saúde evoluam em direção a um modelo centrado no cidadão, no qual um conjunto abrangente de informação do doente deve estar automaticamente disponível para as equipas que lhe prestam cuidados, independentemente de onde foi gerada (local geográfico ou sistema). Este tipo de utilização segura e agregada da informação clínica é posta em causa pela fragmentação generalizada das implementações de sistemas de informação em saúde.

Várias aproximações têm sido propostas para colmatar as limitações decorrentes das chamadas "ilhas de informação" na saúde, desde a centralização total (um sistema único), à utilização de redes descentralizadas de troca de mensagens clínicas.

Neste trabalho, propomos a utilização de uma camada de unificação baseada em serviços, através da federação de fontes de informação heterogéneas. Este agregador de informação clínica fornece a base necessária para desenvolver aplicações com uma lógica regional, que demostrámos com a implementação de um sistema de registo de saúde eletrónico virtual. Ao contrário dos métodos baseados em mensagens clínicas ponto-a-ponto, populares na integração de sistemas em saúde, desenvolvemos um *middleware* segundo os padrões de arquitetura J2EE, no qual a informação federada é expressa como um modelo de objetos, acessível através de interfaces de programação.

A arquitetura proposta foi instanciada na Rede Telemática de Saúde, uma plataforma instalada na região de Aveiro que liga oito instituições parceiras (dois hospitais e seis centros de saúde), cobrindo ~350.000 cidadãos, utilizada por ~350 profissionais registados e que permite acesso a mais de 19.000.000 de episódios.

Para além da plataforma colaborativa regional para a saúde (RTSys), introduzimos uma segunda linha de investigação, procurando fazer a ponte entre as redes para a prestação de cuidados e as redes para a computação científica. Neste segundo cenário, propomos a utilização dos modelos de computação *Grid* para viabilizar a utilização e integração massiva de informação biomédica. A arquitetura proposta (não implementada) permite o acesso a infraestruturas de e-Ciência existentes para criar repositórios de informação clínica para aplicações em saúde.

keywords

Health information technology; regional health information networks; Electronic Health Record; systems integration; Grid computing; HealthGrid.

abstract

Modern health information technology is not just a supporting instrument to good information management but a strategic requirement to provide more efficient and safer health care. Health information technology is a cornerstone to build the future patient-centric health care systems in which a comprehensive set of patient data will be available to the relevant care teams, in spite of where (system or service point) it was generated. Such secure and efficient use of clinical data is challenged by the existing fragmentation of health information systems implementation.

Several approaches have been proposed to address the limitations of the so called “information silos” in healthcare, ranging from full centralization (a single system) to full-decentralized clinical message exchange networks.

In this work we advocate the use of a service-based unification layer, by federating distributed heterogeneous information sources. This clinical information hub provides the basis to build regional-level applications, which we have demonstrated by implementing a virtual Electronic Health Record system. Unlike the message-driven, point-to-point approaches popular in health care systems integration, we developed a middleware layer, using J2EE architectural patterns, in which the common information is represented as an object model, accessible through programming interfaces.

The proposed architecture was instantiated in the *Rede Telemática da Saúde* network, a platform deployed in the region of Aveiro connecting eight partner institutions (two hospitals and six primary care units), covering ~ 350,000 citizens, indexing information on more than 19,000,000 episodes of care and used by ~350 registered professionals.

In addition to the regional health information collaborative platform (RTSys), we introduce a second line of research towards bridging the care networks and the science networks. In the later scenario, we propose the use of Grid computing to enable the massive use and integration of biomedical information. The proposed architecture (not implemented) enables to access existing e-Science infrastructures to create clinical information repositories for health applications.

Contents

List of figures

List of tables

List of abbreviations

Working definitions for the concepts marked with (*) are provided in the next section.

API Application Programming Interface.

BING Brain Imaging Network Grid project (not to be confused with the Microsoft Bing search engine).

CEN *Comité Européen de Normalisation* (European Committee for Standardization)

CNPD *Comissão Nacional de Proteção de Dados* (the Portuguese National Data Protection Agency)

CPOE Computerized physician order entry.

EC European Commission of the European Union.

EGI European Grid Infrastructure.

EHR Electronic Health Record. (*)

EMI European Middleware Initiative.

GERES-med Grid-enabled repositories for eScience medical applications project.

GP General Practitioner (used in this work as the role of the Primary care physician that follows a person/family).

HDA Hospital Distrital de Águeda.

HIETA® Healthcare Integration Engine for Telematic Applications, the middleware component of RTSys implementing the information hub abstraction. HIETA is a registered trademark of the University of Aveiro.

HIEx Health Information Exchange. (*)

HIP Hospital Infante D. Pedro.

HL7 Health Leven 7 (http://www.hl7.org), an ANSI-accredited standards developing organization for health information technology.

HP Portal The Health care Professionals Portal of RTSys.

HP Health care professional.

ICEHR Integrated Care Electronic Health Record. (*)

ICT Information and Communication Technology.

IEETA *Instituto de Engenharia Eletrónica e Telemática de Aveiro* (www.ieeta.pt), the research unit providing the material conditions for the present work.

IHE Integrated Healthcare Enterprise, an international association (with regional chapters) sponsored by Healthcare Information and Management Systems Society (HIMSS) and the Radiological Society of North America (RSNA).

IS Information System.

ISO International Organization for Standardization.

MDS Minimal Data Set (on the clinical context of a Patient).

OGSA Open Grid Services Architecture.

PCU Primary Care Unit.

R-EHR Regional Electronic Health Record, *i.e.* an EHR contributed by multiple organizations in a community/region.

RHIN Regional Health Information Network. (*)

RHIO Regional Health Information Organization. (*)

RIS *Rede Informática da Saúde* (private networking infrastructure for the Portuguese health system).

RTS Project The *Rede Telemática de Saúde* (health telematic network) project (presented in section 1.2).

RTS The deployment of the RTSys system in the region of Aveiro for the purpose of supporting a RHIN. We'll use RTS Project to designate the project, RTS for the deployed instance, and RTSys for the computational system.

RTSys The system (*i.e.* telematic platform) proposed in this work, including the integration middleware (HIETA), access services and portals. In other others: the overall solution. RTSys is the system; RTS is the RTSys instantiation in the region of Aveiro.

US The United States of America.

VO Virtual Organization.

Definition of key terms

e-Health. The application of information and communication technology to health care. For a comprehensive review of the uses of this term see the work of Oh (Oh *et al.*, 2005). Since the scope of application is quite vast and the term is often used as a buzzword, in this work we will rather prefer the expression Health Information Technology, to focus on the 'informatics' component, resorting to e-Health when we mean the broader application context.

Electronic Health Record (EHR). In the words of ISO-TR-20514, the EHR is 'a repository of information regarding the health status of a subject of care in computer processable form, stored and transmitted securely, and accessible by multiple authorized users. It has a standardized or commonly agreed logical information model which is independent of EHR systems. Its primary purpose is the support of continuing, efficient and quality integrated health care and it contains information which is retrospective, concurrent, and prospective' (ISO-TC215, 2005). This concept is refined in ISO-18308, which maintains these terms but considers that the EHR for an individual might not be restricted to a single repository, rather one or more repositories, physically of virtual integrated, that might be scattered across multiple (discrete or interconnected) clinical systems and repositories, each keeping relevant health data fragments in the life of a person (ISO-TC215, 2011).

Electronic Health Record system (EHR system). A system for recording, retrieving and manipulating information in electronic health records (ISO-TC215, 2005).

Health Information Exchange (HIEx). A set of standards, technologies and governance framework put in place to electronically exchange data between health information systems at different health care organizations. The scope of the information being exchange can vary; common examples are medical imaging for diagnosis, ePrescription or summary care records.

Healthgrid. The HealthGrid white paper defines this concept as 'Grid infrastructures comprising applications, services or middleware components that deal with the specific problems arising in the processing of biomedical data. Resources in HealthGrids are databases, computing power, medical expertise and even medical devices' (Breton *et al.*, 2005).

Integrated Care EHR (ICEHR). An Electronic Health Record that supports sharing between different clinical disciplines, different applications connected to the same EHR node,

and between multiple EHR nodes (meaning different EHR systems or locations). The concept is defined in ISO-TR-20514 as a sharable EHR that 'its primary purpose is the support of continuing, efficient and quality integrated health care' (ISO-TC215, 2005). The ICEHR can be seen as a full realization of the EHR concept, including the ability to coherently use information from multiple EHR repositories.

Regional Health Information Organization (RHIO). An organization that oversees and governs the exchange of health-related data among organizations within a defined geographic area for the propose of improving health and care in that community (adapted from (NAfHIT, 2008)).

1 Introduction

The European health care systems are often presented as a major social breakthrough, ensuring that populations have access to high quality and affordable care (OECD-HD, 2011). In Portugal, the national health system is seen as one of the most relevant achievements in the last three decades (Simões, 2010), generally ratted as very good in international assessments (WHO-ROE, 2010).

The maintenance and evolution of modern, universal health care systems represent a tremendous challenge, given the level of investment required and the social expectations on care delivery (EC, 2007c). It is widely recognized that health care systems are under transformation, much pressured by financial sustainability concerns and the need to address new challenges. The new requirements are the result of a changing reality: new mobility patterns of patients and professionals; new expectations from the citizens willing to play a more active role in their own health management; an increasing care demand by an aging population; the ever growing information processing needs to ensure efficient care processes in a digital interconnected world.

In this context, health information technology plays a key role on health care systems sustainability and transformation and holds the potential to deliver improvements at multiple levels (Chaudhry *et al.*, 2006; Institute of Medicine, 2001; Stead *et al.*, 2009; EC, 2007b):

- Evolve from a bureaucratic, institution-oriented system to a patient centric one. Putting the patient in the centre means that it is up to the care organizations, service points and professionals to ensure the required level of coordination to offer an integrated service to the citizen (Shaller, 2007). In such vision, the information of a patient must flow and be accessible to the relevant care providers, anywhere.
- Evolve towards better chronic diseases management and better support of the elderly needs. The development of assisted environments, contributing to gradually replace (a part of) acute care episodes in Hospitals with homecare, is seen as a key evolution to face the population ageing in European countries (EC, 2008b).
- Evolve towards personalized medicine in which the combining of multiple information sources, such as the EHR, bioinformatics databases and environmental data, could enable new individualized therapeutics. Such vision cannot be achieved without the development of new biomedical knowledge models and computational infrastructures (Martin-Sanchez *et al.*, 2004).

Health information technology is an opportunity and a required tool to raise the efficiency of the health care system (OECD-HPS, 2010). Examples of cost reductions and overall process optimization are available concerning, for example, the use of Electronic Health Records (Chaudhry *et al.*, 2006), ePrescription (Dobrev *et al.*, 2010) or the positive impact in the system from involving the patient (Pagliari *et al.*, 2007).

Health information technology is a key to evolve from an health system centred on costs contention to one oriented by clinical results and value-adding for the patient and the society (Porter *et al.*, 2006). Moreover, it is increasingly recognized that health information systems play a crucial role to assure patient safety (Aspden *et al.*, 2004).

The perception of the strategic value of health information technology led the European Union member states to commit to an ambitious eHealth implementation plan, calling for action in the development of infrastructures, interoperability initiatives and legal frameworks (EC, 2004). In the US, the American Recovery and Reinvest Act of 2009 includes substantial financial incentives to stimulate the use of health information technology, specially the adoption of Electronic Health Records (Blumenthal, 2009). Health information technology is in the political agenda.

But there are also obstacles to the introduction and generalization of health information technology, with both success stories and huge failures available (OECD-HPS, 2010). A key problem is the fragmentation of the market and the subsequent existence of a myriad of non-interoperable solutions. Paradoxically, the fragmentation is extensible to the use of standards, questioning their very purpose (Empirica, 2008). The lack of interoperability between health information systems hinders the seamless access to patient data and blocks systems integration, negatively impacting the development of novel services (Stroetmann *et al.*, 2009; Lenz *et al.*, 2007) and the reduction of costs based on scale economies (EC, 2007a; Empirica, 2008). In addition, the lack of coherence in the use of clinical information can originate omissions and potentiate errors (Kohn *et al.*, 2000).

The barriers to the seamless flow of clinical information raised by information systems heterogeneity in health care and the methods to overcome them are at the centre of the present work; Figure 1.1 presents an overview of the topics addressed.

Figure 1.1: Thesis overview as a world cloud. The size of the letters reflects the frequency of each term in the text.

1.1 Health information technology for connected care

Modern health care is a teamwork engaging several professional groups (*e.g.* physicians, nurses, technicians), different medical specialties, and distributed service points (Barros *et al.*, 2007). The intense specialization characterizing the field had consequently originated the proliferation of information and communication technology (ICT) systems and the coexistence of multiple solutions to collect, analyse and preserve clinical information, indispensible for the daily operations.

The progressive digitalization of health information, essential for the innovation development, is also originating a problem of proliferation of systems. The coexistence of heterogeneous systems raises important barriers, which, in part, are opportunities for the computational methods disciplines (to enhance data and processes integration). We can find multiple ICT solutions devoted to access the disperse information pertaining to a patient within an organization (Cruz-Correia *et al.*, 2005), a region (Mäenpää *et al.*, 2009), a country (EC, 2009) and even to address cross-border scenarios (epSOS, 2011). Several national authorities are building health information infrastructures to allow clinical data to flow across regions and health services (Deutsch *et al.*, 2010; Blumenthal, 2009), although these are more likely to be subsets than complete electronic health records (Greenhalgh *et al.*, 2010b). In this sense, the European Union is starting a landmark cross-border, large-scale pilot on patient data exchange, in the scope of the epSOS project (epSOS, 2011).

We will use the term 'connected care' to designate a shared use of the patient data between different professionals, working at autonomous care organizations, for the purpose of safer and more efficient care delivery (EC, 2006). The connected care realization must rely on some electronic information infrastructure to articulate heterogeneous systems and enable (clinical) information to flow between service points (Cruz-Correia *et al.*, 2007).

Looking at the Portuguese context, if a patient receives care from multiple service points at different institutions, chances are that none or only a small part of the related information is shared between care providers. When the information exchange occurs, it is in general unidirectional (from the referring physician to the solicited service point), using tangible media. In general, routine electronic health data exchange is not supported (Figure 1.2).

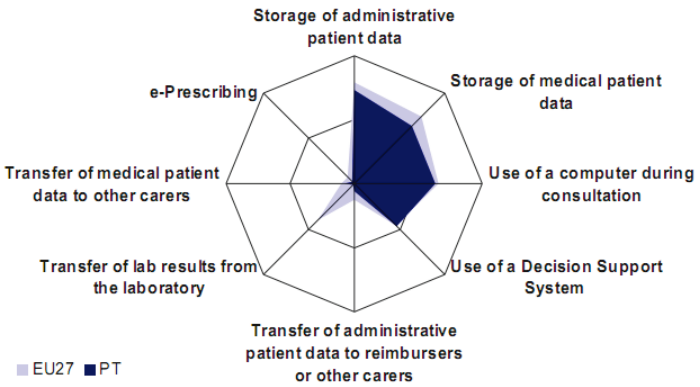


Figure 1.2: Global perspective on the use of eHealth in Portugal (available from (Dobrev et al., 2008)).

This contrasts with the great deal of attention devoted to the introduction of EHR to support the needs of Hospitals, during the nineties (van de Velde, 1992). These investments in health information technology aimed at solving institutional and, often, departmental needs, and the same level of effort was not committed to health information technology to articulate different care providers (Figure 1.2). This state of the technology is often known as 'information islands' in health care (Knaup *et al.*, 2007; Lenz *et al.*, 2007). Although other European countries present a better record on electronic patient data transfer and access, there are still several hard research problems to be tackled (EC, 2007b), such as the semantic interoperability of health information systems, the secure yet ubiquitous access to patient data by the authorized professionals, and the definition of best practices for patient data sharing and their insertion in the organizations (Figure 1.3).

The problem of interconnecting systems in cooperative workflows is generally labelled as system interoperability. When specific methods are needed to ensure that systems and humans involved in information sharing have the same interpretation of data, we are addressing a semantic interoperability scenario (Lenz *et al.*, 2007). The full connected care vision is somewhere at the end of the semantic interoperability implementation roadmap (Stroetmann *et al.*, 2009).

Since it is not possible (nor desirable, to our eyes) to build a single system to address all the possible requirements of health information management (Berg, 2004), the defensible answer is the general adoption of collaborative practices, best served by the use of standards, in which each system is a component in an ecosystem of interoperable parts. Although the domain stakeholders are well aware of the impact and benefits of standards (US-GAO, 1993), the path to standardization in health informatics is slow and difficult to implement, and their use is far from being generalized (Anderson, 2007; Empirica, 2008).

Given such comprehensive challenges in health information technology, what can be done today to transform a system based on 'disconnected institutions' into one centred on the citizen? From a system that communicates insufficiently into one based on sustained sharing? From a fragmented health IT solutions space into one 'ecosystem' of systems? We argue that a possible answer is the adoption of sharing models in the regional health care value chain.

The regional scale provides a more agile context for the introduction of pilots and has a strong correspondence to the referral and transfer patterns of patients. The introduction of

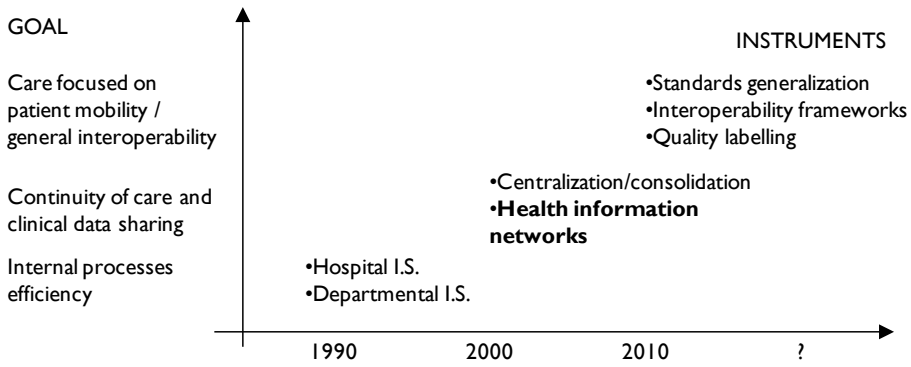


Figure 1.3: A possible summary on major trends in health information technology investments. Health information networks are the main topic of this work.

telematic solutions at regional level (1) immediately facilitates the existing processes of patient referral between service points, (2) enables clinical information sharing in a controlled environment, (3) requires less investments than broader approaches (though still not easy to raise). The bottom-up introduction of scalable architectures in the implementation of Regional Health Information Networks (RHIN) can provide the basis for flexible exploitation models, enabling partner institutions to collaborate in care provision today, by sharing practices and information, in pursue of connected care (EC, 2006).

This line of thought was at the genesis of the *Rede Telemática da Saúde* research project, which designed and implemented a RHIN in the region of Aveiro (the project is presented in the next section). The *Rede Telemática da Saúde* (RTS) network was driven by the local health care institutions (Hospitals and primary care units) seeking to collaborate and share best practices. As a bottom-up effort with limited funding, partners in RTS looked at practical solutions to enable regional collaboration, while keeping ICT architectural styles that could scale to include new partners and use cases.

The principal contribution of this work is the definition and implementation of the RTSys, the telematic platform enabling the regional network. RTSys introduces a middleware solution that can be deployed in an *ad hoc* community of care institutions, to aggregate disperse clinical data. RTSys does not introduce redundancy with respect to the existing systems, nor replaces them. Instead, it provides a virtual health record by federating information in the existing sources. The virtual record is accessible at the point of care and presents information produced and stored at several partner institutions, in a coherent user interface.

1.2 The 'Rede Telemática da Saúde' project

The present work has been started in the scope of the *Rede Telemática da Saúde* (RTS) project and the technical framework we propose (the RTSys solution) was firstly deployed as a result of the project. The RTS project had other outcomes than the telematic platform and, on the other hand, the development of RTSys continued as a research activity after the project has been completed. We can look at the RTS project as the source of requirements and a first instantiation of our work, providing an organizational context and domain users for pilot testing.

The RTS project (2004-2006) supported by competitive funding under the CidadesDigitais programme of the Portuguese Government, aimed at bringing to the Portuguese context regional connectivity experiences that already proved their value elsewhere (examples available in section 2.3.4). The principal promoting partner was Hospital Infante D. Pedro (HIP), the larger hospital in the region of Aveiro. Besides HIP, the consortium integrated the second largest Hospital, Hospital Distrital de Águeda (HDA), and nine primary care units, represented by the umbrella organization Sub-Região de Saúde de Aveiro (six installed and used the solution, Figure 1.4). The technical definition and implementation of the telematic methods were with the University of Aveiro, also a member of the consortium, which has decided to further enhance the resulting system on its own after the funded period.

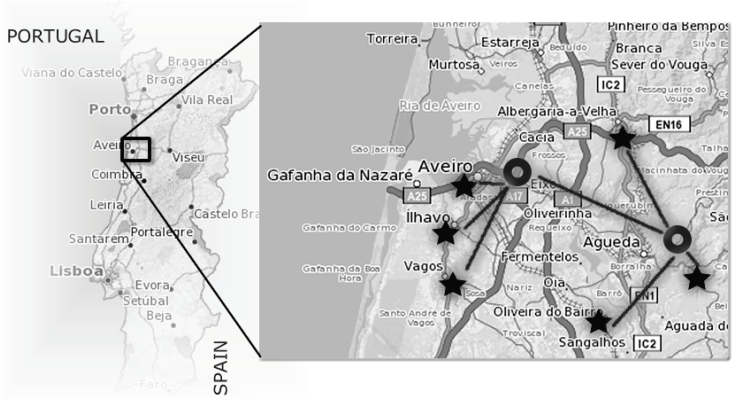


Figure 1.4: Primary care centers (stars) and Hospitals (circles) using RTS. Lines represent existing referral schemes in the region of Aveiro.

The most visible outcome of the RTS project is a portal for the health care professionals, allowing them to access a subset of the Electronic Health Record, autonomously built by aggregating information scattered in multiple information systems in the partner institutions. A complementary outcome is a portal for the citizen to monitor his health agenda (populated with region-wide events) and interact with the partner institutions.

The RTS project provided the source of requirements for the computational system presented in this work. Multidisciplinary teams, involving 42 health professionals and the University of Aveiro research group, worked on the requirements definition. The same clinical institutions received the developed system and collaborated on the pilot use, to which ~200 users received specific training on the Health Professionals portal.

RTSys is not limited to the RTS deployment: we propose a generic and scalable telematic platform to connect regional health care that can be deployed in other scenarios, further stretching the proposed architecture.

Given the scope of the processes supported in RTS and the remarkable complexity of the health care domain as an application area, it would not be possible to develop the RTSys without the openness and fruitful collaboration of the institutions in the RTS consortium and the professionals involved, to whom we are deeply grateful.

1.3 Virtual communities in biomedical e-Science

In many fields of modern science we witness a data deluge and the consequent need for researchers to use computer methods to create models, manage data and run analysis tasks (Hey *et al.*, 2009). The use of computer methods as a tool to enable scientific research is known as e-Science (Hey *et al.*, 2003). This new way of doing science calls for action on the research agencies, which are promoting the use of advanced computing infrastructures (or e-Infrastructures) as an essential technology to structure research communities and solve large scientific problems, taking advantage of the ability to aggregate distributed capabilities

(computing resources, storage space, information sources, scientific instruments) and make them usable in a controlled cyber-environment (SIENA, 2012).

One specific technology enabling the deployment of e-Infrastructures is Grid computing. It provides a distributed computing and data management technology around the concept of Virtual Organizations (VO) of researchers sharing similar requirements (Foster *et al.*, 2001). Grid computing shares similarities with the now popular Cloud-computing, in the sense they both enable a seamless use of scalable computational resources and resort to similar Internet technologies (Foster *et al.*, 2008). Grids and Clouds are complementary rather than opponents (SIENA, 2012), with the first being more focused on the use of advanced computing infrastructures available in research centres to support specific thematic VOs (Foster *et al.*, 2008).

The application of Grid computing to life sciences can be generically labelled as HealthGrid. The HealthGrid vision advocates the use of Grid infrastructures in medical research, health care and life sciences, to enable advanced models of the person and the disease and run complex simulations and analysis (EC-SHARE, 2008). A complete HealthGrid is a Grid able to manage, relate and process information at different scales, from molecule to man, as elaborated in the HealthGrid White paper (Breton *et al.*, 2005), to which our group has contributed.

These new e-Infrastructures raise opportunities concerning the archiving and sharing of medical images and clinical reports for scientific applications, for example, to manage banks of medical imaging cases (Blanquer *et al.*, 2010). New search strategies based on the content of the images (extraction of features), which are very computational demanding, can now be supported in wide-scale infrastructures (Montagnat *et al.*, 2008a).

Grids can also present value to the clinical practice, given that privacy issues are secured. Grids can be used, for example, to share large amounts of data within a virtual community, to provide valuable analysis services to members (*e.g.* computed added detection), or to create long-term cases repositories (Bellotti *et al.*, 2007; Power *et al.*, 2005). The use of Grid infrastructures, however, still exposes much of the underlying systems idiosyncrasies and usually requires users to be familiar with system operations and even the structure of the Grid (Bird *et al.*, 2009).

In this work, we present a second line of research, proposing the use of Grid computing to support the analysis of clinical information originated in care settings. To this end, we describe an architecture to bridge the 'clinical information infrastructures' and the e-Infrastructures. This roadmap can be instantiated to connect RTSys with existing research e-Infrastructures. Although not implemented, several aspects of the overall solution have been prototyped (as discussed in Chapter 7).

1.4 Research goals

The main research goal of this work concerns the definition and implementation of RTSys, a computational environment to enable the coherent access to patient information available in distributed, heterogeneous systems, across different organization domains. This is done under specific constraints: the new system must not disrupt existing information systems; data

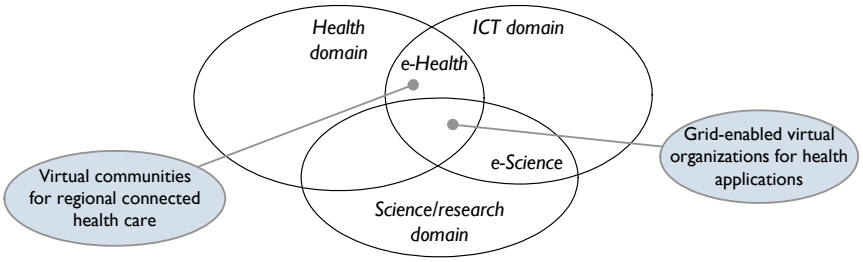


Figure 1.5: Conceptual positioning of the research areas (shaded shapes).

integration should be seamless to the end user and delivered at the point of care; the system must provide a secure environment, observing the regulation context of the health domain.

The coherent and practical access to disparate clinical information sources has been addressed in the context of previous projects within our research group, including multimedia departmental integrated systems (Oliveira *et al.*, 2000), data formats for biosignals interoperability (Cunha *et al.*, 1997) and process redesign for best teamwork in Hospitals (Cruz *et al.*, 2002). The need for 'trans-institutional' architectural styles in healthcare information systems (Haux, 2006), mapping the domain teamwork in electronic supported collaborations, has been a study topic in the group that lead to the RTS project proposal.

The research objectives of this work are related to the design and development of computer methods that, in the context of the Portuguese reality, answer three fundamental problems in systems integration for connected care: an architecture for multi-institution shared data access based on discovery processes; a security model to configure and impose access policies; a services interface for external applications development leveraging on the RTSys processes.

We argue that the introduction of a domain specific middleware, acting as an information hub, can provide the functional abstractions to enable a community view upon the fragmented information infrastructure, alleviating the lack of interoperability and standards. Upon this layer of unification, services with a regional focus can be developed. In this sense, we elaborate on possible use cases for a regional health information network powered by our middleware, with special focus on the definition of the regional, virtual Electronic Health Record.

Besides the theme of health care information integration, we address a second line of work, towards the secondary use of health information in research (Figure 1.5). This research goal focuses on the definition of a computational architecture to enable information available in clinical infrastructures to be used in Grid-enabled infrastructures by virtual research organizations.

1.5 Thesis contributions

In this work, we present the RTSys system, a telematic platform for care institutions to collaborate and share patient information. The RTSys comprises two layers: (1) a flexible health information integration middleware (named HIETA), and (2) access services. The access services

include an API for applications development and two portals, one tailored for the health professionals and another to the citizen.

RTSys has been implemented using free software technologies (no licensing costs for the partner organizations) and deployed in the region of Aveiro, in the context of the RTS project. The Health Professionals portal is now available at eight care institutions. This telematic platform for health care is unique in Portugal and a novel effort concerning the opening of institutional walls to share clinical data with partners for the continuity of care; it has been awarded a National innovation prize[1] and presented in several forums as a relevant advance in the Portuguese eHealth landscape.

With respect to the second line of research, we present a Grid-enabled architecture for using health data in research infrastructure from a conceptual point of view, and partial exploratory implementations are discussed.

The models and approached developed in the context of this work have motivated the publications listed in Table 1.1.

Table 1.1: Outlook of the author's contributions (reverse chronological order).

Publication	**Main contribution(s)**
I. C. Oliveira, N. Silva, I. Veiga, and J. P. S. Cunha, "Supporting nursing care assessment protocols with smartphones," in *International Conference on Health Informatics (HEALTHINF 2012)*, Vilamoura, Portugal, 2012, pp. 81-86.	Presents a system to enable several teams in a region to participate in the monitoring of skin lesions, using network and camera enabled smartphones. The RTS Professionals portal is used for visualization of shared cases.
I. C. Oliveira and J. P. S. Cunha, "Integration Services to Enable Regional Shared Electronic Health Records," in *User Centred Networked Health Care - Proceedings of Medical Informatics Europe 2011*, Oslo, Norway, 2011, pp. 310-314.	Discusses the architecture of RTSys and the outcomes from the deployment in Aveiro.
I. C. Oliveira, E. Dias, L. Alves, J. Barros, J. A. Silva, M. P. Monteiro, A. S. Pereira, J. M. Fernandes, A. Campilho, and J. P. S. Cunha, "Extending a desktop endoscopic capsule video analysis tool used by doctors with advanced computing resources," in *5th Iberian Grid Infrastructure Conference (IberGrid 2011)*, Santander (Spain), 2011, pp. pp. 156-168.	Presents a practical approach to access Grid infrastructures from a clinical desktop to process endoscopic capsule exams.
I. C. Oliveira, L. Alves, E. Dias, D. Pacheco, S. Lima, J. Barros, M. P. Monteiro, J. A. Silva, J. M. Fernandes, J. P. S. Cunha, and A. S. Pereira, "Automated endoscopic capsule analysis using a Grid computing environment," in *Ibergrid: 4th Iberian Grid Infrastructure Conference Proceedings*, 2010, pp. 319-330.	Presents an algorithm and a set of validation experiments to run a topographic segmentation method on a Grid infrastructure for video exams from endoscopic capsule modalities.
I. C. Oliveira, J. P. S. Cunha, D. Pacheco, J. M. Fernandes, M. Pedrosa, L. Alves, and A. S. Pereira, "A neuroscience Grid-enabled portal for the Portuguese Brain Imaging Network," presented at the MICCAI-Grid Workshop, London, UK, 2009.	Proposes the use of a portal system, in which the neuroscience research workflows are mapped, to allow research to define and share data sets and have them analysed in Grid infrastructures.

[1] Hospital of the Future award ("Hospital do Futuro") awarded in 2006 under the category "Integration".

Publication	Main contribution(s)
I. C. Oliveira, J. M. Fernandes, L. Alves, A. S. Pereira, and J. P. S. Cunha, "GERES-med: An Architecture for Grid-Enabled scientific REpositorieS for medical applications," in *Ibergrid: 2nd Iberian Grid Infrastructure Conference Proceedings*, 2008, pp. 163-173.	Proposes a system architecture to build on existing Grid infrastructures to deploy repositories of medical images to be shared in Virtual Communities, specifically to address the requirements of the cardiology and gastroenterology communities.
I. C. Oliveira, J. L. Oliveira, J. P. Sanchez, V. López-Alonso, F. Martin-Sanchez, V. Maojo, and A. Sousa Pereira, "Grid requirements for the integration of biomedical information resources for health applications," *Methods of Information in Medicine,* vol. 44, pp. 161-167, 2005.	Requirements analysis from the perspective of the integrated use of biomedical and health information on the Grid. Actual information databases are presented and possible directions drawn.
Related work, as co-author:	**Main contribution(s)**
J. P. S. Cunha, J. M. Fernandes, I. C. Oliveira, M. Pedrosa, L. Alves, and A. S. Pereira, "The Portuguese BING Network: Towards a Brain Imaging Grid Virtual Community," in *Ibergrid: 3rd Iberian Grid Infrastructure Conference Proceedings*, 2009, pp. 96-105.	Describes the technical deployment of the Portuguese Brain Imaging Network Grid and the capabilities available for the BING community. I.Oliveira contributed to the technical design.
D. Pacheco, I. Oliveira, J. M. Fernandes, and J. P. S. Cunha, "MAGI: A Medical Application Grid Interfacing portal for eScience," in *Ibergrid: 3rd Iberian Grid Infrastructure Conference Proceedings*, 2009, pp. 31-42.	Presents a friendly web portal to run common medical imaging analysis tacks on a Grid infrastructure. This publication received a **best student paper award**. I.Oliveira was an adviser in this work.
V. Santos, D. Oliveira, I. C. Oliveira, and J. P. S. Cunha, "A monitoring toolkit for a distributed clinical data integration engine," in *Healthinf 2009: Proceedings of the International Conference on Health Informatics*, Setubal, 2009, pp. 300-305.	Presents a set of supporting components to monitor the operations of the RTS network. I.Oliveira coordinated the overall implementation.
R. Andrade, I. C. Oliveira, J. M. Fernandes, and J. P. Cunha, "A Grid framework for non-linear brain fMRI analysis," *Studies in health technology and informatics,* vol. 126, pp. 299-305, 2007.	Presents a method to run non-linear, non-parametric analysis of fMRI volumes using Grid computing. I.Oliveira participated in the system design.
J. P. S. Cunha, I. Cruz, I. C. Oliveira, A. S. Pereira, C. T. Costa, A. M. Oliveira, and A. Pereira, "The RTS Project: Promoting secure and effective clinical telematic communication within the Aveiro region," presented at the eHealth 2006 High Level Conference, Malaga, Spain, 2006.	The RTS project concept and goals, proposing a regional health information network for Aveiro. I.Oliveira contributed to the technical design of the system.
V. Breton, K. Dean, T. Solomonides, I. Blanquer, V. Hernandez, E. Medico, N. Maglaveras, S. Benkner, G. Lonsdale, S. Lloyd, K. Hassan, R. McClatchey, S. Miguet, J. Montagnat, X. Pennec, W. De Neve, C. De Wagter, G. Heeren, L. Maigne, K. Nozaki, M. Taillet, H. Bilofsky, R. Ziegler, M. Hoffman, C. Jones, M. Cannataro, P. Veltri, G. Aloisio, S. Fiore, M. Mirto, I. Chouvarda, V. Koutkias, A. Malousi, V. Lopez, I. Oliveira, J. P. Sanchez, F. Martin-Sanchez, G. De Moor, B. Claerhout, and J. A. Herveg, "The Healthgrid White Paper," Studies in health technology and informatics, vol. 112, pp. 249-321, 2005.	The HealthGrid white paper was a landmark report aiming at defining the field and proposing possible applications and research areas. I.Oliveira was the editor of Chapter 5 in this white paper, entitled "Genomic Medicine Grid".

1.6 Thesis structure

A schematic way to look at the thesis content is supported by the roadmap from Table 1.2. In detail, the present text is structured as follows:

In Chapter 2, we review the background concepts and the state of the art, focusing on the principles to represent patient health information and the enabling technologies to share clinical data. A brief overview of the current situation in Portugal concerning these topics is included.

In Chapter 3, we specify the use cases and requirements for a regional collaborative health information platform. We depart from the tangible reality of the region of Aveiro and the RTS project to set the system functional scope.

In Chapter 4, we elaborate on the design of the computational methods and system architecture required to support the selected use cases. A middleware for health information integration (named HIETA) is proposed as an enabling component to fulfil the virtual, regional Electronic Health Record abstraction.

In Chapter 5, we discuss the system implementation, detailing the interactions that occur (collaboration between system components) and the technologies deployed.

In Chapter 6, we present and discuss the results attained, concerning the proposed system design and its pilot use in the region of Aveiro.

Chapter 7 is an extension of the main research line, elaborating on a possible architecture to bridge RTSys and existing Grid infrastructures for health applications. Since this topic is contained and different from the main research line, we chose to keep it in its own chapter, including the sections dealing with the state of the art, proposed models and results.

To model the information systems entities, we use the Unified Modelling Language notation (Booch *et al.*, 2005) whenever appropriate.

Table 1.2: A possible roadmap for eHealth in support of personalized healthcare and related topics in this work.

Principles for health systems (re)structuring	**Technical requirements (contributions needed from health informatics)** [a)]	**Related topics addressed in this work**
Patient centric care (the system revolves around the patient)	Technical, syntactic and semantic interoperability between health information systems. Sharing models and practices agreed by the care providers. Collaborative working models enabled by information hubs *and workflow systems*.	Use cases for the RTSys (Chapter 3) Computational model for a regional information hub (Chapter 4) and its implementation and deployment (Chapter 1)
Patient involvement (the patient is entitled to information and active participation)	Information hubs to enable a single, coherent interface to the patient. Security models (authentication and authorization) Patient-oriented friendly portals. *Integration of patient-generated content and automatic health monitoring.*	
Adapt therapies to individual health characteristics and conditions	Methods for the massive aggregation and usage of biomedical information (databases, storage resources, computing elements) Repositories of clinical modalities (*e.g.* medical imaging) for research.	Requirements and an architecture for health applications on the Grid (Chapter 7)

[a)] Topics in italics are not explored in this work and are presented just for reference.

2 Review of enabling technologies and state of the art

The use of information and communication technology and electronics in health care is commonly designated eHealth (often written as e-Health). Although an accepted neologism, it does not imply a single, universal definition (Oh *et al.*, 2005), rather addresses a wide range of applications (a possible categorization of the application areas is provided in Table 2.1, according to vision of the industry). In this work, we prefer the use of Health Information Technology to denote the use of digital data and computer methods for health applications, including, for example, the use of Electronic Health Records (EHR), decision support systems, electronic physician order entry, health information networks, medical imaging and patient-oriented devices and homecare solutions. We are specifically addressing the 3rd group of application in Table 2.1, 'health information networks and distributed electronic health record systems'.

In this chapter, we present background concepts and the state of the art with respect to the use of Electronic Health Records in networked environments and selected enabling technologies.

2.1 The Electronic Health Record

Healthcare is an information and knowledge intensive domain and it has long been recognized that traditional paper-based approaches cannot keep the pace with the innovation in health systems (Dick *et al.*, 1997). The use of information technology plays a key role to enhance the efficiency and safety of health care systems (Institute of Medicine, 2001; Stead *et al.*, 2009; EC, 2007b).

At the heart of the health information technology is the Electronic Health Record (EHR), the structured and lifelong repository of a patient health status and health care (Tang *et al.*, 2006b; Wyatt, 1994). It is an essential technology to enable new computer-based methods in medicine towards data automation and safer practice (Dick *et al.*, 1997). The concept, however, is used with much latitude; see, for example, the work of Häyrinen for a comprehensive review (Häyrinen *et al.*, 2008).

2.1.1 PURPOSE AND CONTENT OF THE ELECTRONIC HEALTH RECORD

The medical record of a patient tells the story of his health conditions, the assessment performed by health care professionals and the care provided over several encounters. It provides a shared information repository essential to enable the collaboration between multidisciplinary

Table 2.1: Market view on eHealth application areas (adapted from (EC, 2007a)).

Category	Description and Examples
I. Clinical information systems	Specialized tools for health professionals within care institutions such as Radiology Information Systems, Nursing Information Systems, Medical Imaging, Computer Assisted Diagnosis, Surgery Training and Planning Systems. Tools for primary care and/or for outside the care institutions such as general practitioner and pharmacy information systems.
II. Telemedicine and homecare, personalised health systems and services	Services for disease management, remote patient monitoring, tele-consultation, tele-care, tele-monitoring, and tele-radiology.
III. Health information networks and distributed electronic health record systems	Regional, national and cross-borders integrated health information networks and supporting services such as e-prescription, e-referrals and federated patient data access.
IV. Secondary usage non-clinical systems	Systems for health education and health promotion of patients/citizens such as health portals or online health information services. Specialised systems for researchers and public health data collection and analysis such as bio-statistical programs for infectious diseases, drug development, and outcomes analysis. Support systems such as supply chain management, scheduling systems, enterprise resource planning, which support clinical processes but are not used directly by patients or healthcare professionals.

teams involved in an on-going episode and across different moments in time. Collectively, the medical records provide a unique source of medical knowledge, from which it is possible to conduct public health studies, learn and research (Stroetmann *et al.*, 2009; PWC, 2009) (Figure 2.1).

The highly influential study conducted by the Institute of Medicine in the United States, in the late nineties, on the opportunity to move towards electronic records (Dick *et al.*, 1997) concluded that the use of Computer-based Patient Records (CPR) should be the standard way for health care professionals and organizations to capture medical and other patient related data. The study argues that CPR is an essential technology to enable the reform of the health care system to cope with upcoming patterns in patient information management: the increasing mobility of citizen calls for transferable information; an aging population will raise the care demand and the amount of information to be managed; improved levels of efficiency in health systems need to build on data automation; medical science advances need to access extensive amounts of research data; and growing awareness on patient safety concerns call for more demanding data

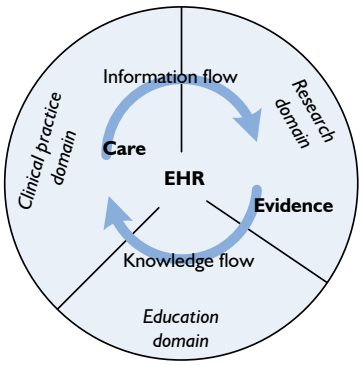


Figure 2.1: The pivot role of the Electronic Health Record in the flow of information and knowledge in health care (adapted from (Chen, 2009)).

Table 2.2: Common designations found in literature related to the EHR concept.

EHR main characteristics	Specific focus or constraints	Related designations
Digital repository of a person's health data; Securely shared between authorized users; Primarily to support continuing, collaborative health care; With secondary uses in health system governance, research and education.	Limited to or mainly focused on the needs of medical specialities.	Electronic Medical Record (EMR)
	Focus on acute care episodes, especially with respect to Hospital settings.	Electronic Patient Record (EPR)
	Partially under the control of the Patient, with some information self inserted.	Personal Health Records (PHR).
	Assembled on demand, by federating multiple health record systems.	Virtual EHR.
	Mainly used in the US in a wide range of senses, varying from the EMR to the EHR (but falling into disuse).	Computerised Patient Record (CPR) or Computer-based Patient Record.

management standards. The conclusions are drawn for the United States but easily transferred to other realities. These early orientations are confirmed by more recent work with the Institute of Medicine, *e.g.* (Aspden *et al.*, 2004), and in several actions promoted by the European Commission, *e.g.* (EC, 2004; EC, 2008a).

The understanding of the core EHR functions evolved and left a trail of different but related definitions that can be found in the literature, summarized in Table 2.2 (and reviewed in (Häyrinen *et al.*, 2008)). The use of the 'health' term rather the once preferred 'clinical' and 'medical' denotes the change on the focus from the care process to the patient, from the acute care to wellbeing, from a internal documentation asset to a collaborative space contributed by different disciplines, spawning from birth to death.

For the European Union, the EHR is 'a comprehensive medical record or similar documentation of the past and present physical and mental state of health of an individual in electronic form, and providing for ready availability of these data for medical treatment and other closely related purposes' (EC, 2008a). ISO defines the basic Electronic Health Record as 'a repository of information regarding the health status of a subject of care in computer processable form' (ISO-TC215, 2005). The same technical report notes different levels of EHR content sharing to conclude that when different clinical disciplines can share the EHR information between different EHR systems and sites, then we are in the presence of the Integrated Care EHR (IC-EHR); it is a EHR that 'has a standardised or commonly agreed logical information model which is independent of EHR systems. Its primary purpose is the support of continuing, efficient and quality integrated health care'. The IC-EHR concept abridges not only information on the disorders, but potentially all health events.

In a recent report to the Office of the National Coordinator for Health Information Technology (USA), the EHR definition explicitly requires interoperability support (NAfHIT, 2008): 'an electronic record of health-related information on an individual that conforms to nationally recognized interoperability standards and that can be created, managed, and consulted by authorized clinicians and staff across more than one health care organization'.

The key characteristics of the ISO IC-EHR definition are summarized in the first column of Table 2.2 and collectively describe the EHR concept as we interpret it in this work.

Purpose of the EHR

Getting the EHR right is challenging (Häyrinen *et al.*, 2008). The challenge starts with the difficulty to answer what is the right EHR and the information structure that meets the extensive set of quality and functional requirements (ISO-TC215, 2011; HL7/ISO-TC215, 2009; Hoerbst *et al.*, 2010a).

While the primary purpose is to address medical care use cases (*e.g.* structured documentation and communication of patient's conditions, repository of clinical documents and results, etc.), it also provides the basis for legal evidence, research and clinical governance. A current view on the purposes of EHRs is available from the British Royal College of General Practitioners guidelines for electronic records best practices (DH+RCGP+BMA, 2011) and summarized in Table 2.3.

What is in the Electronic Health Record?

The ISO's definition presents the EHR as a repository that should have a standardized information structure (or at least agreed) and independent from the concrete EHR systems implementation (ISO-TC215, 2005). In practice, we observe a great heterogeneity in the EHR structure and content (Häyrinen *et al.*, 2008), though there are several standards providing concrete information models to the content of the EHR (Kalra, 2006; Eichelberg *et al.*, 2005). Examples for the content model are available from the Health Level 7 Reference Information Model (HL7, 2012b), ISO-13606-Part 1 in conjunction with ISO-13606-Part 3 (ISO-TC215, 2008), and ISO-12967-2 (ISO-TC215, 2009b).

Functional requirements for EHR systems

The EHR is implemented in EHR systems which provide the required building blocks, tools and interfaces to manage the patient data (Tang *et al.*, 2006b). Jha *et al* provide a list of requirements for a comprehensive EHR system for Hospitals (Jha *et al.*, 2009b), covering the needs of multiple clinical units (Table 2.4).

The main reference internationally accepted to the set of functions that the EHR should meet is the ISO-18308 standard (released in 2004 and updated in 2011); it provides a technology-

Table 2.3: Summary of the EHR purposes (adapted from (DH+RCGP+BMA, 2011)).

Purpose	**Applications**
Primary: clinical.	Support the care of individual patients. Assist in the clinical care of the practice population.
Primary: non-clinical.	To meet administrative, legal, and contractual obligations.
Additional (secondary uses)	Feed cognitive-systems (e.g.: decision support, adverse events detection). Clinical governance, professional development, education and training. Commissioning and healthcare planning.
Emerging	Health records created in one health environment are increasingly likely to be accessed for viewing and/or editing in other health environments. Patients to have increasing control over their health records.

independent specification of the EHR system architecture requirements (but not a concrete architectural design). The requirements are organized in five groups: business, clinical information representation, communication and interoperability, ethical and legal, and confidentiality (ISO-TC215, 2011).

In a complementary effort, Health Level 7 (HL7) developed a standard for the functions of an EHR system, the HL7 EHR System Functional Model (HL7/ISO-TC215, 2009). This specification does not convey implementation technologies, nor implies any clinical messaging system, but sets a comprehensive list of functions and conformance criteria to delimit and compare the functionality of EHR systems. The new release (Release 2, close to final balloting) defines 322 functions and 2310 conformance criteria checks.

2.1.2 THE IMPACT AND ADOPTION OF EHR SYSTEMS

The EHR has been put in the centre of health information technology and praised so highly that is often seen as its 'holy grail'. The introduction of EHR systems is expected to:

- Introduce process optimization, leveraging in the digital representation of clinical information (*e.g.* results from tests available sooner, anywhere), surpassing the limited paper-based approaches. Reduction in documentation time, however, is not likely to be achieved or significant (Poissant *et al.*, 2005).
- Reduce the risk of errors by making the patient information more accessible for informed decision making (EC, 2007b).
- Reduce costs related to duplication of diagnostic procedures (Jha *et al.*, 2009a).
- If shared, enables a better coordination for the continuity of care (RCGP-HIG, 2009).
- Secondary uses of EHR are usually acknowledged as instruments for research (Elger *et al.*, 2010), governance (PWC, 2009) and public health (Kukafka *et al.*, 2007).

Although the potential benefits are ambitious, it is difficult to establish the evidence of benefits and successful adoption. Uslu *et al* have concluded that there is fair evidence that costs

Table 2.4: Functional requirements of an EHR system (adapted from (Jha et al., 2009b)).

Class:	**Clinical documentation**	**Test and imaging results**	**Computerized order entry**	**Decision support**
Required functions / areas:	Demographic characteristics of patients Physicians' notes Nursing assessments Problem lists Medication lists Discharge summaries Advanced directives	Laboratory reports Radiologic reports Radiologic images Diagnostic-test results Diagnostic-test images Consultant reports	Laboratory tests Radiologic tests Medications Consultation requests Nursing orders	Clinical guidelines Clinical reminders Drug-allergy alerts Drug–drug interaction alerts Drug–laboratory interaction alerts Drug-dose support

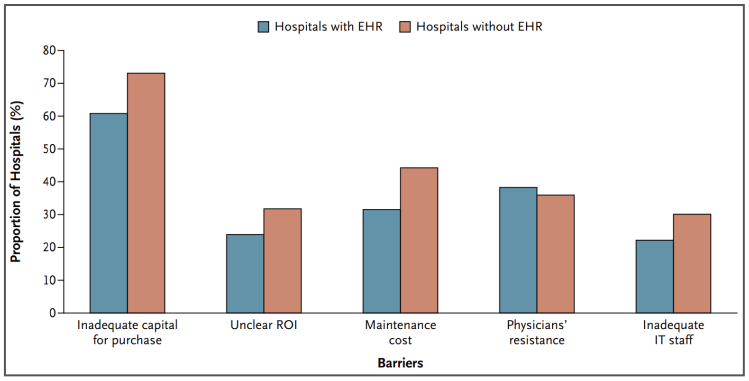


Figure 2.2: Barriers to the introduction of EHR systems in the United States (available from (Jha et al., 2009b)).

can be reduced, but the treatment quality improvement is not clear (Uslu *et al.*, 2008). In a recent report to the European Commission, Dobrev *et al* show that EHR systems are not quick wins; they have very long return on investments and require net cash injections; eventually, the associated long-term socio-economic gains to society make them worthy investments (Dobrev *et al.*, 2009).

In the US, the adoption of EHR is very low, either by Hospitals (Jha *et al.*, 2009b) and by doctors (Blumenthal, 2009), which led to a recent major financial push to stimulate their introduction (Blumenthal *et al.*, 2010). The limited penetration of EHR systems has been related to several barriers (Figure 2.2); the perception of hindrances slightly differs in hospitals with and without EHR systems.

We can conclude that the full potential of EHRs has not yet been achieved (Häyrinen *et al.*, 2008) and additional R&D is necessary, especially with respect to the integrated usage of different systems (Garde *et al.*, 2007; Knaup *et al.*, 2007; Bisbal *et al.*, 2011).

2.2 Sharing for the continuity of care

Moving from the individual sphere of information analysis to public communication with an audience involves an all new set of concerns. It requires a medium for the message to reach its target (interfacing technique) but also that, once received, the message is understood, which often implies that the parties share a language (structure) and set of concepts (representations of the world). This familiar image can help us to grasp the substantial differences and challenges when going from the EHR as an intra-organizational tool into a 'system of EHR systems' approach, aiming at the aggregate use of patient information.

2.2.1 HEALTH INFORMATION EXCHANGE

Digital health information exchange (HIEx) aims at improving the effectiveness of the health care systems, by feeding the care processes directly with information available in different organizational domains (Kaelber *et al.*, 2007). The theme has gained much attention with the

recent stimulus pack from US government[2], which, in practice, sets information exchange a requirement for organizations to be eligible for the financial support (Edwards *et al.*, 2010). HIEx plays also a central role in many other national strategies for health IT (Mäenpää *et al.*, 2009; Progress-Consulting *et al.*, 2011).

HIEx is expected to improve (1) patient safety and (2) the efficiency of the health system, leading to cost reductions (Walker *et al.*, 2005). HIEx can improve patient safety by making more information available at decision time and applying data checks (*e.g.* medication, laboratory, radiology, public health, etc.) in the communication among providers (Kaelber *et al.*, 2007). The comprehensive work from Chaudhry *et al* found evidence of quality and efficiency improvements, but is uncertain with respect to the generalization of results to other settings (Chaudhry *et al.*, 2006). Findings from Frisse *et al* suggest that for a regional scale HIEx may positively impact the decrease in laboratory and radiology procedures, fewer admissions and lower overall costs with emergency care (Frisse *et al.*, 2007). These are consistent with the increased efficiency of outpatient care reported by Mäenpää *et al* (Mäenpää *et al.*, 2011).

Despite the strategic advantages of HIEx, it still is in a maturation process; more research is needed to fully understand its value and implications (Edwards *et al.*, 2010). Fontaine *et al* argue that the positive impact of HIEx needs further empiric evidence (Fontaine *et al.*, 2010). In particular, HIEx is of less value in the absence of a clear regulation framework to contextualize the collaboration between (autonomous) organizations (Doosselaere *et al.*, 2008).

The recent developments in the US, in which HIEx is addressed as a key part of the healthcare system reform (Adler-Milstein *et al.*, 2011b), also lunched the discussion about the 'meaningful use' of information across organizations. Professionals and hospitals are expected to implement EHRs but they need to talk to each other, in the sense that the sharing is useful fand relevant for the patient treatment process, and participants can safely use the information from other organizations, *i.e.* the meaning is preserved (Blumenthal *et al.*, 2010).

There are different models to implement health information exchange (Cruz-Correia *et al.*, 2007). We can distinguish between two main paradigms: a communicational approach, based on messages exchange, and an architectural approach, based on some middleware layer. Message exchange systems are, by far, the most prevalent; they use a clinical messaging standard, specially the suite of protocols from HL7 (HL7, 2012a), to implement interfacing mechanisms between discrete systems. The messaging concepts may be used as the interfacing solution in a higher level approach, as in the IHE Integration Profiles (IHE), detailing not only the data structures but also the actors and interactions occurring in the exchange scenarios.

The middleware strategy defines a common information model and offers services to applications for shared information access. Usually, the data is not replicated (no central repositories) but exposed as a single virtual source. A middleware approach will typically resort to an enterprise-level decoupling technology, such as Web Services (Mykkänen *et al.*, 2007).

[2] Specifically: American Recovery and Reinvestment Act of 2009, Title XIII - Health Information Technology, Subtitle B—Incentives for the Use of Health Information Technology, Section 3013, State Grants to Promote Health Information Technology.

Table 2.5: Characteristics of different deployments of EHR systems (available from (ISO-TC215, 2005)).

	Local-EHR system	**Shared-EHR system**	**EHR Directory Service**
Scope and purpose	Individual local health providers	Local care communities Regional or national	National Transnational
Type of EHR	Non-shareable EHR ICEHR	ICEHR	Index to ICEHR
Type of data	Detailed local data	Shared data	Meta-data index
Granularity of data	Fine	Local care community or extended community (regional/national)	N/A
Contributors and access to EHR	Local health providers	Local care community or extended community (regional/national)	N/A
Custodian/maintainer	Health Care Facility (Hospital, GP clinic, etc)	Local health authorities, GP custodian, etc.	Public health departments or similar

These two approaches, further analyzed when considering alternative methods to build regional health information networks in section 2.3, provide an high-level view of foundation options when defining a sharing environment, and a very rough simplification of the technical and semantic design issues involved (Lenz *et al.*, 2007; Edwards *et al.*, 2010).

2.2.2 SHARED ELECTRONIC HEALTH RECORDS

The ISO-20514 distinguishes three deployment models for the EHR: local, shared and directory (Table 2.5) (ISO-TC215, 2005). The local deployment targets the detailed care record kept autonomously by care organizations, and would contain all relevant information for the patient context known at that organization. The shared model concerns the use of a subset of the EHR between multiple entities for the continuity of care, in a secure way. The directory lookup model addresses a more decoupled relationship, such as the cross-border exchange of clinical data. In this work, we are aiming at information technology to enable the use of a shared subset with a community focus.

The potential benefits of shared records inherit from the general advantages presented for HIEx which include improvements in the safety of care delivery, process efficiency and cost effectiveness; nonetheless, the clinical evidence available in favour of shared EHR deployment is still weak (for a comprehensive discussion see (RCGP-HIG, 2009)). We can expect long-term benefits from EHR systems that are designed to interoperate, but not early return on investments (Dobrev *et al.*, 2010).

Not all EHRs were conceived for sharing and not all fit this purpose, though modern EHR systems should adopt architectures that enable to use the EHR in regional and global initiatives (Haux, 2006; ISO-TC215, 2011). The use of shared EHR is a paradigm shift that introduces new requirements and challenges (RCGP-HIG, 2009), for example, how to combine the pervasive access and ease of use with secure access (van der Linden *et al.*, 2009); or the patient concerns about his information being shared by multiple organizations (Bergmann *et al.*, 2007; Simon *et al.*,

2009). Shared records can enhance integrated care, but governance and new practices must be careful planned (Featherstone *et al.*, 2012).

The share model does not require the entire EHR to be available. In fact, more often than not the shared information is a subset of the information kept at local EHR, selected for that specific purpose, commonly designated as a patient summary. The patient summary approach is being used in several European countries, either to provide basic information to the care giver in a unexpected encounter scenario, or as a partial strategy to compensate the fragmentation of health information technology (EC, 2006).

In this work, we are addressing the shared model from the perspective of a read-only access to a summary record available in the region, contributed by fragments in the EHR systems at multiple care providers.

2.2.3 SEMANTIC INTEROPERABILITY IN HEALTH INFORMATION TECHNOLOGY

The semantic interoperability is a cross-cutting dimension of health information technology. It addresses the problem of ensuring that computers can exchange information automatically while preserving its meaning, thus enabling the secure reuse of clinical information by heterogeneous systems, multiple professionals and different application contexts (Stroetmann *et al.*, 2009).

Semantic interoperability concept

Automating the information exchange between different systems requires agreeing on the connectivity protocol and interfacing technologies, but also ensuring that the receiver interprets the information with the same meaning as it was produced by the sender, *i.e.* that the participants (human or systems) share a common interpretation of the information (Stroetmann *et al.*, 2009). For the European Commission, semantic interoperability is the capability to ensure that 'the precise meaning of exchanged information is understandable by any other system or application not initially developed for this purpose' (EC, 2008a). This is a demanding (not to say problematic) requirement since it extends the preservation of meaning to systems that eventually were not initially conceived to engage in cooperative health care workflows.

The comprehensive report by Stroetmann *et al* defines semantic interoperability in health information technology as the ability to: (1) exchange, understand and act on the patient data and other health information; (2) across multiple actors, with different linguistic and cultural backgrounds, including health professionals, patients and other actors; (3) within or across the boundaries of health systems, in a collaborative manner (Stroetmann *et al.*, 2009). In this definition, the technology is almost secondary; the fundamental properties of interoperability are in the collaborative outcomes when different people, information systems and health systems can work together. The report proposes four levels of interoperability achievement (Table 2.6). Lopez *et al* also note that interoperability requirements are far more wide-ranging than 'just' information compatibility (Lopez *et al.*, 2009), including business process orientation and appropriate security services, for example.

In the scope of this work, we address the secure sharing of distributed EHR systems that meets at least the level 2a, by adopting an integration middleware. In the current state of the art, Layer 3 is still out of reach and significant research is still needed to pave the way for such deployments (Stroetmann *et al.*, 2009; Lenz *et al.*, 2007).

Impact of semantic interoperability

The lack of interoperability between health information systems and EHR systems in particular is a major barrier to achieve efficiency gains in this domain, as attained for other industries (Grimson *et al.*, 2000; Edwards *et al.*, 2010). Without semantic interoperability, it is not safe to use health information technology in distributed scenarios (EC, 2007b), and, therefore, automation. Failing to understand the risks raised by the lack of interoperability may led the health information technology investments to turn into an 'expensive, digital chaos' (Kadry *et al.*, 2010).

Interoperable EHR systems provide an essential foundation to attain systemic benefits (Dobrev *et al.*, 2009) such as: (1) secure and unambiguous sharing of clinical data between different professionals and heterogeneous systems; (2) adoption of clinical decision support and workflow systems that use several patient data sources; (3) enable the secondary use of EHR data in research, public health and governance; (4) integrate personal health record systems and auto-recorded data in health systems; (5) allow the linking of EHR content with external sources, specially with scientific repositories and citizen oriented web resources. The emergence of new paradigms in health information exchange, such as the Personal Health Records and biomedical knowledge integrative models, only reinforces the need for semantic interoperability.

In 2004, the European Commission issued an Action Plan with the required political roadmap towards an European eHealth area; the plan highlighted the need for the adoption of interoperable EHR systems, based on agreed standards (EC, 2004). Again, in 2008, the Commission produced a sound statement on this topic, recommending to the Member States the adoption of interoperable EHR systems to support cross-border exchange of health data, in order to 'enhance

Table 2.6: A taxonomy for semantic integration levels (adapted from (Stroetmann et al., 2009)).

Level	Short description	Additional details
Level 0	No interoperability.	No technical infrastructure available to ensure connectivity.
Level 1	Technical and syntactical interoperability (but not semantic interoperability)	It is possible to access remote data, but human intervention is required to interpret the content.
Level 2	Partial semantic interoperability: Level 2a: unidirectional semantic interoperability Level 2b: bidirectional semantic interoperability of meaningful fragments	Some shared fragments are encoded and can be automatically used by computer systems with precise meaning.
Level 3	Full semantic interoperability	Technical or syntactic heterogeneity are overcome and systems able to use automatically all the relevant (remote) information

the quality and safety' of patient care in the European space (EC, 2008a). Despite the efforts at the European level to bring coherence to eHealth technologies — which ultimately would improve cross-border care delivery and lead to new service opportunities resulting from the defragmentation of the market (EC, 2007a)— it is still up to each country to organize its own health system and health information technology policies. In this context, the epSOS project is deploying a ground-breaking pilot on cross-border patient data exchange, forcing the level 2b of interoperability (Table 2.6).

Based on a comprehensive literature review, Edwards concludes that additional research is still needed to fully understand the role of interoperability in health information technology and how to overcome the barriers limiting health information exchange (Edwards *et al.*, 2010). A similar conclusion and a comprehensive review of the challenges and possible solutions are drawn in the SemanticHealth project (Stroetmann *et al.*, 2009).

We can conclude that interoperability between EHR systems is not widely available, nor will it be in the short term. Besides a technical and standardization issue, it also calls for leadership and high level decisions on investments and policies (Stroetmann *et al.*, 2009; EC, 2008a). In the words of Berg, 'technology is crucial, but secondary' and the models for organization development need to come first (Berg, 2004).

2.2.4 STANDARDS FOR INTEROPERABLE EHR SYSTEMS

The interoperability between EHR systems is not yet solved. Nevertheless, significant contributions have been made towards the harmonization of the EHR systems' requirements, their content and logical architecture, resulting in the existence of multiple standards with complementary purposes (Eichelberg *et al.*, 2005). The use of standards in EHR systems implementation is expected to deliver a foundation to avoid the vendor lock-in phenomena (Blind, 2004), structure this much fragmented market (EC, 2007a) and, at some level, enable a future-proof EHR (Beale, 2002), in the sense it would fit in an unpredictable future ecosystem of systems. Several reviews on EHR related standards are available in the literature (Eichelberg *et al.*, 2005; Kalra, 2006; 2008) and also in the state of the art report produced by the think-tank (formed in 2009) in charge of outlining a strategy for a national EHR system in Portugal (RSE-WG, 2009b).

Outlook of health information technology standards

Several standards have been proposed to guide the development of EHR systems (Kalra, 2006). These standards target different aspects of EHR systems implementation, from high-level functional requirements, domain knowledge capturing, information structures, up to the level of value sets for field coding (Empirica, 2008).

In this topic, we should note the work of accredited international standards organizations, such as ISO Technical Committee ISO/TC 215, CEN Technical Committee CEN/TC 251 and Health Level 7. Some standards produced under CEN activities or HL7 are also recognized and published by ISO, originating dual standards. Other high-impact industry-driven standards exist, especially DICOM (Mildenberger *et al.*, 2002) for medical image exchange (by the Diagnostic Imaging and

Therapy Systems Division of the US National Electrical Manufacturers Association), and IHE technical frameworks to facilitate the adoption of existing medical informatics standards for cross-enterprise information exchange (IHE). A different approach is found in the openEHR initiative, which provides a comprehensive set of open specifications for EHR systems and open reference implementations of some components (Garde *et al.*, 2007). The European Commission has recently sponsored a comprehensive review study of the ICT standards available for the health area (Empirica, 2008).

A criticism to the state of the art in health informatics standards availability is rather than being too few, they seem to be too many (Table 2.7). We can observe the existence of competing standards for the same purpose (for example, ISO-13606 parts 1 and 3, and HL7 RIM for information content). The transformability between models for the same purpose may be possible, but is not ensured without the loss of accuracy (McCay *et al.*, 2008). We will now briefly present the most relevant to the present work.

The two-level approach

A common problem with standards is that they focus on syntactical and structural normalization of information models, but fail to provide a framework to explain the meaning of the data (semantics) in ways that computer methods can use to support knowledge-enabled tools (Beale, 2002). The semantics are conveyed in the written specification and part of the usefulness of the standard depends on a conservative use of the models. The users (of the standard) cannot take advantage of its definitions to construct dialects to accommodate specific needs.

A different approach is found in archetype-based standards (*e.g.* openEHR and ISO-13606), adopting a two-level modelling approach (Bisbal *et al.*, 2011). The two separate levels are *information* and *knowledge*:

- Level 1: a basic immutable set of concepts and relationships forms the Reference Model of the universe of discourse. This model is sparse and does not define the actual information structures to be implemented.
- Level 2: an archetype system, based on a well-defined language for templates specification, allows defining the valid information constructs for a given purpose, using the vocabulary from the Reference Model. The archetypes impose several types of constraints on the intended data set (*e.g.* multiplicity, associations between concepts) and support bindings of data values to terminologies.

The two-level modelling approach for capturing the health domain knowledge provides more flexibility by separating informational and knowledge concepts (Beale, 2002), but is still little disseminated (Wollersheim *et al.*, 2009). While bringing powerful concepts to increase the semantic expressiveness of health data, it also poses implementation difficulties: legacy EHR systems cannot immediately benefit from it; bindings to external terminologies must be defined; and, probably the more thorny issue to address, archetypes still need to be shared in controled and consensual ways, calling, therefore, for some sort of archetypes governance to avoid a profusion of incompatible dialects (Garde *et al.*, 2007). One should note that in order for an archetype-based approach to succeed, it still needs to be ground in some consensus that makes

the templates sharable, future-proof and useful. In addition, both openEHR and the ISO-13606 multipart standard target the representation of information for EHR systems, but not the surrounding ecosystems in which they need to operate.

Selected ISO/CEN standards

ISO-TR-20514 (Electronic health record – Definition, scope, and context) provides a set of definitions relevant to understand and contextualize the several uses of the EHR concept (ISO-TC215, 2005). It provides a basic classification and details the purpose and characteristics of the EHR and EHR systems. EHR can be shared or isolated; when shared, the report identifies different levels of integration: between disciplines; between different applications at one site; between different EHR nodes, forming the Integrated Care EHR.

ISO-TS-18308 (Requirements for an electronic health record architecture) defines high-level requirements for systems processing and transmitting EHR information (ISO-TC215, 2011).

Table 2.7: Standards relevant for the development and sharing of EHR (grouped by origin and purpose).

ISO/CEN Standards	**HL7 Standards**	**openEHR**
Requirements:		
ISO 18308:2011 Health informatics -- Requirements for an electronic health record architecture (2004, rev. 2011) EN/ISO 12967-1:2009 Health informatics -- Service architecture -- Part 1: Enterprise viewpoint (2009) EN 13940-1 - Health informatics - System of concepts to support continuity of care - Part 1: Basic concepts (2007)	Electronic Health Record-System Functional Model, Release 1.1 (ISO/HL7 10781:2009)	
Information content:		
ISO 12967-2:2009 Health informatics -- Service architecture -- Part 2: Information viewpoint (2009) ISO 13606-1:2008 Health informatics -- Electronic health record communication -- Part 1: Reference model ISO 13606-2:2008 Health informatics -- Electronic health record communication -- Part 2: Archetype interchange specification ISO 21090:2011 Health informatics -- Harmonized data types for information interchange EN/ISO 12967-1:2009 Health informatics -- Service architecture -- Part 2: Information viewpoint (2009)	HL7 Clinical Document Architecture (Release 2: 2005) HL7 Clinical Statements	openEHR EHR Information Model (2008) openEHR Archetypes
Architecture and interfacing		
ISO 13606-5:2008 Health informatics -- Electronic health record communication -- Part 5: Interface specification EN/ISO 12967-1:2009 Health informatics -- Service architecture -- Part 3: Computational viewpoint	HL7 Clinical Document Architecture (Release 2: 2005) HL7 Clinical Statements HL7 SOA Retrieve, Locate and update Services DSTU	
Security		
ISO 13606-4:2009 Health informatics - Electronic health record communication - Part 4: Security ISO/TS 22600-1:2006 Health informatics -- Privilege management and access control		

Systems that meet these requirements are expected to fit health care delivery and to be 'clinically valid and reliable'. While discussing the required characteristics, the standard does not convey technical specifications for their implementation. This standard proposes a refined vision of the EHR to include 'one or more repositories, physically or virtually integrated', thus abridging the sharing and exchange of EHR information. The specification structures the requirements in complementary axes: business, information representation, interoperability, ethical and legal issues, and privacy protection.

ISO-13606 (Electronic health record communication) is a multi-part standard consisting of complementary specifications to enable the exchange of EHR information: a Reference Model of the universe of discourse; an Archetype system to develop concrete archetypes; a set of terms lists for attributes in the Reference Model and informative Reference Archetypes; a set of Security requirements; and the interfacing protocols and payload to request and provide EHR content (ISO-TC215, 2008). This multipart standard was first published by CEN (with the same number) and is now accredited in both organizations. The standard is focused on ensuring a secure and semantic-rich communication of parts (or the all) of one individual's EHR; it does not imply any specific implementation technology. This standard has received much attention due to the use a two-level modelling approach, which makes it essentially compatible with the OpenEHR initiative (Garde *et al.*, 2007).

Another example of a CEN standard being harmonized with ISO is the ISO/EN-12967 (Health informatics Service architecture). This multipart standard builds on the results from several European research projects to define a reference architecture for health information systems, advocating the use of a middleware to federate heterogeneous systems, with well-defined healthcare common services (ISO-TC215, 2009a).

Health Level 7

Health Level Seven (HL7) is both an organization and the resulting collection of standards to normalize electronic data exchange in health care environments. Initially focused on the United States health system and the need to optimize the insurance industry business processes, it has managed to deliver the most successful international clinical messaging specification, with HL7 version 2 being supported by many vendors worldwide. The adoption of HL7 Version 2 showed inconsistencies among implementations and, as a major improvement, Version 3 provides a formal Reference Information Model — RIM (HL7, 2012b), to clarify the information content of messages (which use the concepts defined in the RIM in different arrangements). HL7 CDA (Clinical Document Architecture) defines a document mark-up standard based on XML to write clinical documents (Dolin *et al.*, 2006) using the concepts defined in the Reference Model. A CDA document can be communicated in a HL7 message or used independently. The CDA standard distinguishes three levels of conformance: level 1, the less demanding, only requires a document to provide a valid CDA header (*e.g.* message body can be free, non-interpreted text or binary data); level 2 requires the use of section-level templates (based on the RIM Act class); level 3 requires that all entries are semantically encoded to support full machine-to-machine readability.

Integrating the Healthcare Enterprise (IHE) Profiles

An important sharing framework, with sound results in medical imaging exchange, is the Integrating the Healthcare Enterprise (IHE) initiative (IHE). IHE has developed a set of technical frameworks for different medical areas (*e.g.* Radiology, Laboratories, Cardiology, etc.). Each technical framework provides Integration Profiles that describe the required capabilities to address real-world integration scenarios, in a practical industry-oriented way. The Technical Frameworks use wide accepted industry standards in medical informatics in their implementation (*e.g.* DICOM protocol to retrieve images, HL7 messaging and CDA) and are reviewed and expanded in regular basis. The IHE suite is being used to structure national strategies to integrate disparate EHR systems, for example, in Austria (Schabetsberger *et al.*, 2010) and France (Lagouarde *et al.*, 2007).

The Continuity of Care Record and the Continuity of Care Document

The Continuity of Care Record – CCR (ASTM-E31.25) is a standard in health information technology from ASTM International, suitable to forward the clinical context of a patient from a service point or practitioner to another. The CCR is written in XML according to the specified schema, and includes the major information topics concerning the continuity of care, including demographic and administrative data, allergies, diagnosis, problem lists, medications, lab results and the care plan (not all topics are required to be filled). CCR is not intended to document the full EHR of a patient, rather to include the information most relevant for the continuity of care, resulting from one or several encounters. The CCR has been criticized for falling short as a general purpose health information exchange standard, lacking the adaptability to cover advanced interoperability scenarios (Ferranti *et al.*, 2006).

The CCR can be transformed into an HL7 compliant CDA, which is also written in XML, forming the Continuity of Care Document – CCD (HL7), as jointly specified by HL7 and ASTM. The American National Standards Institute (ANSI) Healthcare Information Technology Standards Panel (HITSP), a partnership in charge of fostering the adoption of health information technology standards in the US, endorsed the use of CCD to communicate extracts of the EHR. The resulting HITSP-C32 specification (HITSP, 2008) further constraints the use of CCD, providing detailed orientations on how it should be used for patient summaries.

Medical terminologies and coding systems

The safe exchange of health information requires not only the parts to agree on the structure of the EHR, but also that they share the interpretation of field-level values. For this purpose, the value sets can be constrained by the use of controlled terminologies. A related (but different) concept is that of an ontology, which defines the semantic relationships between concepts, providing a map to the knowledge of a given domain (Bodenreider *et al.*, 2006).

An example of a widely accepted medical terminology is the Systematized Nomenclature of Medicine – Clinical Terms (SNOMED-CT) by International Health Terminology Standards Development Organization (IHTSDO) (SNOMED, 2012). This terminology provides an extensive

vocabulary to supply the EHR with precise encoded concepts (medical histories, disorders, lab test results, treatments, etc).

Another relevant example is the international standard for the classification of mortality data International Classification of Diseases (ICD) by the World Health Organization. It has been adapted by the US government to best serve the classification of diagnosis, originating the Clinical Modification variant (ICD-CM, 2011). The current version in use is the ICD-10-CM, with some countries still using ICD-9-CM. All public Portuguese hospitals use ICD-9-CM since in practice it is required for the reimbursement procedures.

The International Classification of Primary Care, Second edition (ICPC-2) is used in Primary care practice to describe encounters, classifying clinical activities (*e.g.* medication and procedures) and patient data (*e.g.* symptoms and complaints) (WHO, 2003). The main clinical information systems for GPs in Portugal, the SAM-SINUS used in almost all Primary Care units, supports the use ICPC-2. GPs, however, are not required to classify every encounter and the recording of classifications is in practice limited.

There are multiple terminology systems available and no single terminology (or ontology) encompasses all the needs of the health domain (see, for example, the reference text provided by (Hammond *et al.*, 2006)). In fact, we can anticipate that multiple terminologies will keep being used and evolving (McDonald *et al.*, 1998), challenging the implementation of information systems.

2.3 Regional Health Information Networks concepts and examples

There are multiple instances of Regional Health Information Networks worldwide, with different scope and objectives. In this section, we characterize a set of features that regional networks for care are expected to deliver and discuss common models used in their implementation.

2.3.1 THE CASE FOR REGIONAL HEALTH INFORMATION NETWORKS

A Regional Health Information Network (RHIN) is a telematic platform that allows the secure access to remote data sets kept at several health care organizations, shared to support collaborative care (Tsiknakis *et al.*, 2002). A related concept is that of a Regional Health Information Organization (RHIO), which is a mission-driven organization focused on enhancing the quality, safety and efficiency of health care within a confined geographic area (NAfHIT, 2008). Yet another term is Regional Health Care Network, which can be seen as a merger between the previous two (Oates *et al.*, 2000). In this work, we are especially concerned with the ICT methods to enable the collaboration and less attention will be devoted to the organizational models; we will prefer the term RHIN thus bringing the focus to the telematic platform.

The scope of the RHIN varies (Progress-Consulting *et al.*, 2011): there may be large regions, each running a specific and more or less autonomous health system under its own health

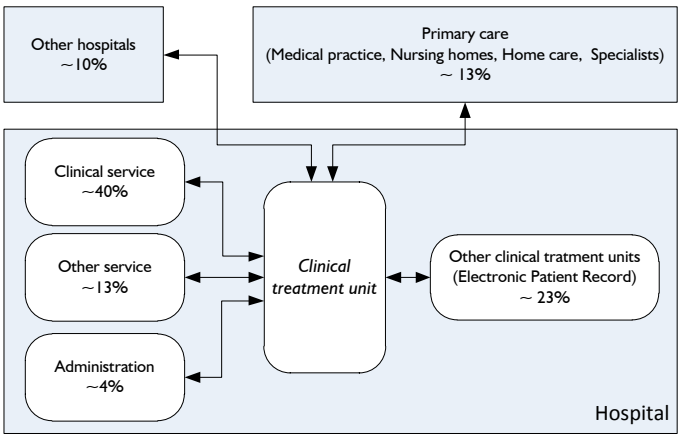


Figure 2.3: Communication flows to support a clinical treatment unit (adapted from the PICNIC project dissemination materials, for the Danish health system).

authority (*e.g.* Spain and Italy), or smaller regions, closer to the metropolitan scope, which are willing to connect, for example, primary and secondary care organizations, such as in the RTS project (presented in section 1.2).

Despite the scale, we are looking for a set of fundamental characteristics that characterize a health information network: (1) different domains/organizations autonomously run clinical information systems; (2) there is some legal entity (or some form of consortium) in charge of governing the network; (3) the partner organizations engage in the use of health information technology solutions to support cooperative workflows, sharing information and best practices.

Each RHIN will define its own goals and use cases to meet the needs of the health system in which it operates. However, all RHINs will have to address the problem of sharing clinical information in secure and controlled ways, available from distributed, heterogeneous systems (Mäenpää *et al.*, 2009; Lenz *et al.*, 2007). The value and impact of the RHIN depends on the ability to enhance information quality and availability (Kripalani *et al.*, 2007), by delivering content proceeding from multiple systems, recorded in different moments in time, by different vendors, with different purposes and scope (Oates *et al.*, 2000).

Constructing a RHIN is a system integration problem (Grimson *et al.*, 2000), thus requiring the development of computational methods typical from distributed systems: how to ensure that the relevant (scattered) data is readily available where and when it is needed; how to identify, authenticate and authorize domain actors; how to coherently identify patients across institutions; how to interface with remote, heterogeneous systems; how to conciliate disparate information models. Systems integration *per se* is not trivial (Sheth *et al.*, 1990), which is only aggravated by the aforementioned requirements of the health domain. We should, therefore, start by reflecting weather the value of RHIN is worth the effort.

Regional connectivity closely follows the health care system structure, in which multiple care organizations cooperate following a set of referral rules within a region. In other words, health care is regional (Oates *et al.*, 2000). Many health information technology initiatives are therefore originated in regional authorities (Progress-Consulting *et al.*, 2011). The need to exchange information effectively occurs inside and across the organization walls (Figure 2.3).

Table 2.8: The main outcomes of RHIN reported in research literature (adapted from (Mäenpää et al., 2009)).

Category	Main outcomes evidenced in literature
Flow of information	Improved access to clinical data and more timely information. Improved clinical data exchange between professionals. A negative factor is the complexity of clinical data exchanged.
Collaboration	Improvement in communication and coordination, resulting in better multidisciplinary teamwork in the region.
Process redesign	Process resigned impacted clinical effectiveness, affecting: improved effectiveness, time saved, supported workflow, supported patient health care plan process, improved decision making and quality of life.
Usability	There are issues with systems complexity, security concerns and less than optimal usability, though operational financial benefits are recognized.
Organization culture	On the positive side, there are commitment to the network attitudes and a sense of participation by regional staff. More sceptical attitudes also exist, with respect to resistance to change and often the lack of strategic vision.

Despite the collaborative nature of care at the regional scale, ICT investments and strategies are typically focused in the organization (Haux, 2006). This is precisely the gap that a RHIN aims to bridge.

Mäenpää *et al* identify four main areas of RHIN outcomes: information flow, collaboration, process redesign and system usability; a supplementary area that may be considered is the organizational culture (Table 2.8). These positive impacts of RHIN can broadly be classified in two main groups: clinical and process efficiency. On the clinical side, the RHIN can improve the sharing of clinical data between health care actors, ultimately contributing to safer decision making (information available for decision making is more comprehensive and up to date); this expectation is supported by the literature (RCGP-HIG, 2009; Stead *et al.*, 2009; Kaelber *et al.*, 2007; EC, 2007b; Bates *et al.*, 2003; EC, 2006). On the process side, financial benefits can be attained by decreasing duplication of services, optimized teamwork (by introducing procedures that can be accomplished online, such as electronic orders entry) and generic optimizations from electronic data processing (OECD-HPS, 2010; Chaudhry *et al.*, 2006; Walker *et al.*, 2005; Dobrev *et al.*, 2009).

Although there is evidence available on the potential financial benefits that can be attained by connected care (Stroetmann *et al.*, 2006; Walker *et al.*, 2005), securing those benefits seems to be challenging and the financial sustainability of RHIO uncertain (Adler-Milstein *et al.*, 2008; Maffei *et al.*, 2009). In the US, many RHIO face financial challenges and present limited activity (Adler-Milstein *et al.*, 2008). This calls for caution when deciding the goals, scope, business model and technical approach to enable regional connectivity (Maffei *et al.*, 2009). In this work, we argue that the most important value of a regional collaboration is centred on the sharing of clinical data to enhance the clinical context of patients (Knaup *et al.*, 2007) and this can be the foundation to a bottom-up endeavour.

2.3.2 ARCHITECTURAL PATTERNS FOR CONNECTED CARE

There is no single best solution to deploy the required ICT to evolve from discrete 'information silos' towards comprehensive IT infrastructures, connecting multiple organizations (Kuhn *et al.*, 2006; Knaup *et al.*, 2007). Instead, multiple architectural approaches and technical

options are available (Cruz-Correia *et al.*, 2007). In this section, we present selected patterns to structure a health information network from the perspective of choices that the enterprise architect should evaluate.

The role of Patient Summaries

A Patient Summary is a reduced version of a person's EHR, scoped for the purpose of the continuity of care; it contains enough information to transmit the essential clinical context of a patient, as defined by domain experts, to support, for example, emergency care or first-time consultation use cases (Remen *et al.*, 2011). The scope of the Patient Summary will naturally vary from implementation to implementation, but typically includes demographic data, allergies and clinical alerts, active diagnosis/problems and chronic or current medication medication (EC, 2006). Additional information domains can be used, such as nursing summaries or oral care. A good example is available from the epSOS project specifications for the purpose of exchanging patient data in cross-border scenarios (epSOS, 2010). The (ASTM-E31.25, 2012) and Continuity of Care Document (HL7) specifications can be used to support the standards-based exchange of patient summaries (though more likely in the US health system).

The Patient Summary approach does not aim at implementing a full-fledged EHR solution but, instead, at defining a common denominator (that can be extracted from EHR systems) to be shared by the relevant actors (Baumlin *et al.*, 2010). This often proves to be a viable balance between the value of sharing clinical data and the difficulty to implement the required organizational and ICT support for interoperability. Several countries are pursing Patient Summary approaches in their national clinical information sharing initiatives; fifteen Member States in the EU reported the existence of national plans to their implementation (2006); similar concerns were reported for the United States, Canada and Australia. The British case is one of the best documented in which much has been invested, although independent evaluations have found that the benefits were less than expected (Greenhalgh *et al.*, 2010a). Patient Summaries need to be complemented with other health IT initiatives (EC, 2004) and are often part of national strategies that also address cross-cutting needs as electronic prescriptions or professionals registries, for example.

Patient Summaries are also an interesting approach to push interoperability in health IT (EC, 2006), supporting sharing practices and models within a constrained scope (this strategy is being adopted in the epSOS project (epSOS, 2010) to force cross-border clinical data exchange). In this sense, they provide a valuable and controlled approach to start building shared practices in a community. As we will detail later, a Patient Summary view is adopted in our solution with a similar goal, not as formal clinical document, but as a metaphor to scope the information to be exchanged.

Centralized and distributed approaches

The shared access to information contributed by multiple actors in the RHIN can rely on a single repository for the clinical data or use integration methods to aggregate content from

multiple sources as needed. We can distinguish between three main deployment models with respect to the implementation of the clinical data repositories: centralized, federated and hybrid.

In a centralized approach, there is a single 'physical' repository, accessed by all partners. The central system may be the reference EHR system for the network (the organizations would not need a local deployment) or a consolidated, materialized view obtained by ETL (extract, transformation and loading) processes from the source EHR systems. Queries are promptly satisfied with the data consolidated in the repository (Halevy, 2001). In a federated model, the organizations retain the ownership of data and all their operational responsibilities. On top of the existing sources, an additional layer provides the required methods to locate and retrieve relevant information on-demand (Sheth *et al.*, 1990). Since information from the source systems is not replicated, queries performance may be affected by the network response and systems operational conditions. A hybrid approach combines aspects from both, with some degree of data replication in the central repository, and federation methods to harvest data just-in-time.

Maro *et al* identify five main practical advantages in a decentralized: (1) institutions retain the control of 'their' data; (2) largely mitigates legal and privacy issues; (3) builds on existing solution and does not required the set up of high-end central data repositories; (4) organizations may choose not to disclosure part of the information; (5) data holders can track and authorize requests for every use of the data they are custodians (Maro *et al.*, 2009). These points, which are more related to the governance of the network than the technical implementation, play a fundamental role to facilitate the acceptance by organizations.

Another strong motivation to use a federated model is that it does not force existing systems to be discontinued or even changed. A federated approach would leverage on the existing capabilities and procedures and provide a new computation layer for integrated access — which is the strategy adopted in our work.

Communicational and architectural approaches

Health information exchange is typically implemented under one of two paradigms: communicational and architectural (Blobel *et al.*, 2006). A communication paradigm is based on the point-to-point exchange of standardized clinical messages and is best illustrated by the successful HL7 standard (HL7, 2012a) (although version 3 already present some aspects that are not just clinical messaging, such as the use of a Reference Information Model to drive message specification). An architectural approach, on the other hand, proposes a reference system architecture, including domain models and well-defined services (Grimson *et al.*, 2000). The ISO 12967 standard (evolving from known as CEN HISA standard) provides an example of such specific architectural layer to interface with the health enterprise information (ISO-TC215, 2009c). The message paradigm supports data integration, but lacks on the functional aspects (Lenz *et al.*, 2007). For a comparison of the principles of the two approaches, see, for example (Blobel *et al.*, 2006).

The two approaches are essentially different, but not incompatible. The use of HL7 messages, in particular, can be used as the content transport solution in an higher-level architectural approach (López *et al.*, 2010). IHE integration profiles, for example, are somewhere in between, in

the sense they build on messages (for data interfacing) but provides guidance to structure the functional integration (IHE, 2012).

2.3.3 FRAMEWORKS FOR HEALTH INFORMATION SHARING

Semantic integration of information and processes (functional integration) in health care is more complex than in other domains, given the fragmentation of implementations and the demanding requirements of this industry (*e.g.* privacy, security, safety). There are, however, frameworks available to construct the capabilities of a RHIN, providing architectures and standards for the underlying ICT methods.

The CEN/ISO-13606 and OpenEHR approach

CEN/ISO-13606 and OpenEHR although not identical, share the same conceptual approach, and it is accepted that they are generically compatible (McCay *et al.*, 2008). Both provide an archetype-based interoperability paradigm and we will discuss them together, referring specifically to OpenEHR for examples and details.

The two-model approach facilitates the development of RHIN by providing the mechanisms that allow domain experts to elaborate their own semantic models for shared clinical information. Domain stakeholders fully retain the ownership of the semantics. At the requirements specification phase, a consensus process should describe the valid information structures to be shared. This is done by specifying archetypes, using the Archetype Definition Language. Archetypes can be registered in a broker system to be shared in a wider community, if appropriate. Terminologies, when used, are bound to data members, in the archetype definition, clarifying the value-sets for the exchanged data.

The openEHR mechanisms support the construction of content models either for a new system-to-be, or to represent existing systems. This content models (archetypes) can be used to establish semantic mappings between heterogeneous models and enable automated transformations (Chen *et al.*, 2009).

The openEHR specifications are freely available and the archetypes mechanism does not imply a conformance test by external entities. The public availability of the specifications and tools make openEHR appealing to use (especially in research projects), but the experience with using it to implement regional health information networks is still limited (Chen *et al.*, 2009).

The Integrated Healthcare Enterprise approach

The most successful industry-led interoperability framework for HIEx is the Integrated Healthcare Enterprise (IHE). IHE is structured in domains (*e.g.* Patient Care Coordination, Radiology, Cardiology, Pharmacy, etc.) and each domain has a technical committee in charge of developing the relevant collection of specifications. The IHE approach to interoperability is based on the concept of Integration Profiles, which describe how accepted standards should be applied in practice to meet the needs of a specific healthcare integration case.

As a foundation principle, IHE aims at reusing established standards, such as HL7 (HL7, 2012a) and DICOM (Mildenberger *et al.*, 2002), in the health domain, and ebXML (OASIS) and Web Services (Erl, 2004), in the IT implementation domain. The benefit is the definition of specific Profiles for concrete integration problems and the validation practices in-place. The introduction of new profiles includes a specification phase (with domain experts and the industry) and a verification phase. The integration tests are performed in large events, the Connectathon, bringing together multiple providers to test implementations against each other.

Each IHE Profile explains, for a given integration problem, which components in a distributed healthcare environment (called Actors) are involved and which information exchange episodes should occur (called Transactions), providing the required guidelines for system implementers. Each Domain is therefore responsible to develop the Profiles relevant to solve integration problems in that area. Profiles are supported by Technical Frameworks specifications, which give the technological details for compatible systems implementation.

The IT Infrastructure domain is a transversal one (others are practice oriented, *e.g.* Cardiology, Laboratory, Radiology, etc.), providing foundation blocks for cross-enterprise use cases. An example is the Cross-Enterprise Document Sharing (XDS) that enables the exchange of clinical documents between health organizations. XDS distinguishes between the Document Repository actor (where the document is persisted) and Document Registry (where the index of existing documents is kept for looking up). Additional actors include Document Sources, Document Consumers and Patient Identity Sources. It is assumed that the participating organizations have agreed to collaborate and share a set of policies and infrastructures, thus constituting an IHE Affinity Domain. A RHIN would be a typical example of an Affinity Domain. The XDS profile relies on messaging standards from OASIS/ebXML (for the system-level encoding) and HL7 (for the clinical messages).

The IHE adoption is progressing rapidly and fits both the needs of large national strategies for health information sharing (Lagouarde *et al.*, 2007; Schabetsberger *et al.*, 2010), as the use cases for regional health information networks (Donnelly *et al.*, 2006). Contributing to its widespread is the abundant support from the industry and the pragmatics of the underlying specifications.

While IHE is very active and its specifications publically available, there are some issues with its implementation. Since IHE interoperability is based on published profiles developed by domain experts in the IHE technical committees, implementers and health organizations depend on the specifications that are approved (though there is a channel to propose changes); it is not feasible for a given community to choose and control, for example, specific structures and semantics for the clinical documents.

Service-Oriented Architectures for interoperability in health information technology

Service-oriented architectures have gained much attention in recent years as a technical and architectural approach to integrate heterogeneous applications in complex enterprise environments (Papazoglou and van den Heuvel 2007). The central concept in this kind of architectures is the use of service interfaces to functions and data. A service is a coarse-grained software entity with a well defined and platform neutral interface description, that is remotely

accessible in the network and performs relevant business functions on business objects. The use of the business term reflects that a service provides an action with visibility and relevance at the business-level perspective (Leymann *et al.*, 2002). Each service is self-contained and loosely coupled (does not depend on the state of other services).

Services, existing in multiple nodes in the network (endpoints), must connect to each other and applications, using an industry standard transport protocol, such as connectivity technologies defined by W3C known as Web Services (W3C). W3C Web Services (WS) use internet-friendly communication protocols and technologies, making extensive use of XML (W3C, 2000) for transport (of parameters and results), service announcement and discovery. A frequent option to implement Web Services makes use of Simple Object Access Protocol (SOAP) (W3C) to transport data instances over HTTP. The platform neutrality of the service interfaces description and the use of common internet technologies make the use of Web Services very attractive to support integration and reuse of software components enterprise-wise and between different organizations.

The Service-Oriented Architecture is an enterprise architecture pattern (Josuttis, 2007) that makes use of:

- Reusable services, able to participate in different business contexts (*e.g.* new business functions). To fit the reusability goal, the service is self-contained and has a well defined responsibility.
- Loosely coupled services. The client (consumer) needs only to be informed of the provider contract (access interface and data structures) and no other details on the internal implementations. Services do not assume any particular usage scenario on the client side to accomplish their functions.
- A deployment infrastructure for connecting services and applications, ensuring remote invocation and transport, usually known as Enterprise Service Bus.
- A managed mechanism for services to collaborate in support business workflows. This includes the announcement and discovery of service contracts and a workflow governance context.

The SOA decoupling and reuse principles are suitable for heterogeneous application integration scenarios, since they allow encapsulating software components and exposing them through standard interfaces. SOA provides the appropriate abstraction level to map technical capabilities and business requirements (Papazoglou *et al.*, 2007).

Almost all complex enterprise-scenarios requiring the integration of heterogeneous software components may benefit from SOA (Josuttis, 2007). Not surprisingly, SOA is increasingly used in healthcare information technology (Katehakis *et al.*, 2007; Yang *et al.*, 2011) and is expected to play a key role in future health information systems (Winter, 2009).

SOA does not replace the existing standards for health information technology. Mykkänen *et al* note that SOA and Web Services have the potential to increase interoperability in healthcare, but the balance between the benefits from services connectivity and the benefits from the health

care semantic and syntax standards is not being generally achieved (Mykkänen *et al.*, 2007). To this end, the Healthcare Services Specification Project, for example, a collaboration between Health Level Seven and the Object Management Group, is working on specifications that blend the SOA business-alignment principles with HL7 wide acceptance for health data encoding (HSSP). Also the IHE Integration profiles are available as Web Services from solution vendors, which can be used to implement common cross-enterprise health care activities (Painter *et al.*, 2012). The SOA, applied as a general enterprise-class interoperability technical and architectural approach, or combined with the semantic and syntactic specificity of health care standards, can bring increased flexibility to the interoperability challenge in health care (HSSP, 2008).

Other frameworks from industry

Microsoft developed a set of guidelines to encourage the development of health information networks called Microsoft Connected Health Framework Architecture and Design Blueprint (Microsoft, 2009). This study describes a transformation process in healthcare using connected systems and open standards. The Connected Health Framework (CHF) proposes the use of the following key strategies: decoupling of components through service orientation; layered data access through federation (keep data as local as possible); federated security delegated in trusted parties to simply identity and roles management.

We can identify in this reference architecture for the health care domain (which is vendor-agnostic and does not bind to any particular technology) a set of components enabling 'connected care' (Figure 2.4): services to handle patient consent, authorization of health professionals,

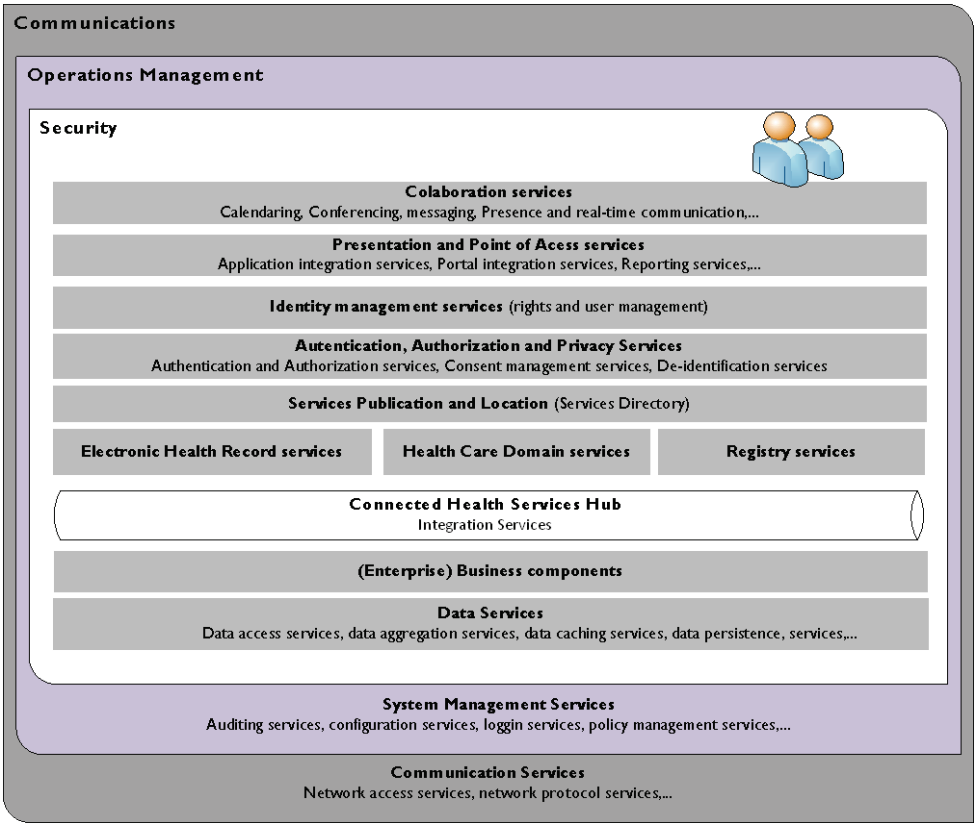


Figure 2.4: An enterprise architecture for connected systems in health care (adapted from (Microsoft, 2009)).

management of the EHR, registry services for federation support, etc. A specific layer in the architecture (Connected Health Services Layer) ensures the functions related to technical data and services integration between autonomous systems. The four-part report provides a comprehensive review and a rich discussion on the technical challenges with respect to deployment of integration technology on the health domain (Microsoft, 2009).

Contrasting with this vendor-agnostic analysis and design blueprint, we can mention the offer from Intel, the 'Intel SOA Expressway for Healthcare', which is presented as a fully-integrated SOA appliance for the health domain (Intel, 2012). The value proposition of this product is based on the performance of the message gateway, the compliance with health information exchange standards (specially the IHE integration profiles and extensive support of HL7 CDA) and integrated security. The natural deployment scenarios for this kind of solution would include the set up of the backbone in a health information network (security, document exchange, patient indexes, etc.) or the implementation of a federated service bus to connect different partner organizations.

2.3.4 SELECTED HEALTH INFORMATION NETWORKS

There are many health information networks implemented or under implementation around the world. Some of them present a regional scope and others aim at embodying a National strategy to enhance clinical information availability. Many state-level and regional initiatives can be found in the United States, in which 197 potential RHIO have been identified (Adler-Milstein *et al.*, 2011a). A more structured approach can be found in Canada, with a coherent national initiative towards interoperable EHR systems running since 2003 (Dennis, 2005). In Europe, the report on national eHealth priorities (EC, 2007c) identifies key strategic projects around the concept of a national EHR system in several countries (Austria, the Czech Republic, Denmark, Estonia, Finland, Romania, Slovakia and Spain) or, alternatively, a shared summary of it (Denmark, Finland, Greece and Italy). In this report, Portugal is lacking on health information sharing, still addressing the setup of the networking infrastructure as a priority.

These initiatives are consistent with the perception that integrated access to patient data originated in different service points may contribute to better care (EC, 2004; Baumlin *et al.*, 2010; Kaelber *et al.*, 2007). The following cases present selected health information networks and how they address the goal of health information sharing.

The PICNIC project: seed research in Europe

Though not a RHIN itself, the PICNIC project (IDT-1999-10345) developed ground-breaking research on the use of ICT to enable regional collaborative care (Saranummi *et al.*, 2005). The results have been validated in 6 regional pilots, over 3 countries. A central argument in PICNIC is that reference components implementing *de facto* and *de jure* standards are key enablers to information systems interoperability in regional health care. The project proposes an architectural approach, based on a set of common components for the health domain and offers an open source reference implementation of the following modules: Messaging (XML messages for referral, examination report and reimbursement), Shared Records Service, Patient

Identification Service, Collaboration Service. The components, which are functionally specified, could be procured using commercial of-the-shelf software. In addition, a conformance test methodology is proposed to ensure that new components can be plugged into the infrastructure. The results of the project influenced the deployment of the Danish health information network and the HYGEIAnet, in Crete. We will use the later to discuss some architectural options.

HYGEIAnet: early regional connectivity in Crete

HYGEIAnet is a RHIN sharing commonalities with our work. It has been implemented in Crete to make available integrated patient records on-line, by incorporating distributed patient data, available at heterogeneous systems (Tsiknakis *et al.*, 2002). Information was kept at the primary sources, and a component-based architecture proposed to federate clinical data. The federation model implies the adoption of a global information schema and mapping of local concepts into the global ones, in which information sources push local information into the shared components.

The implementation of HYGEIAnet incorporated the conceptual approach proposed in the PICNIC project and adopted a set of middleware common components to enable healthcare cross-organization integration (Tsiknakis *et al.*, 2002). These components were initially implemented in CORBA, following the relevant OMG standards (*e.g.* clinical information sources access use the OMG COAS (OMG, 2001)). In later evolutions of the platform, the implementation was revised to support a service-oriented approach based on web services, encapsulating the existing functionality (Katehakis *et al.*, 2007). The major functional components are the following (Figure 2.5): Person Identification Service (PIDS) to consolidate patient identities; Lexicon Query Service (LQS) which manages controlled medical terminologies; Clinical Observation Access Service (COAS) for accessing the contributing information sources; Resource Access Decision (RAD) Facility to support resource-oriented access control policies; I-EHR Indexing Service (I-EHR IS) to index the sources of primary information for performance and scalability; I-EHR Update Broker (I-EHR UB) to detect updates at the information sources; the Health Resource Service (HRS) to

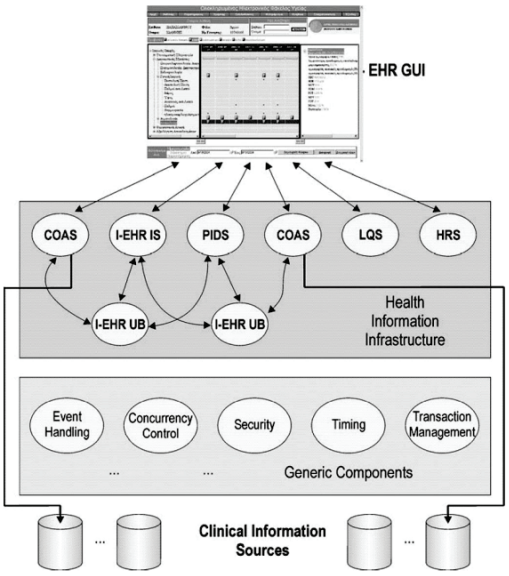


Figure 2.5: HYGEIANet components (available from (Katehakis et al., 2007)).

identify professionals and organizational units. As we will discuss later, similar functional components, with different names and packaging, were developed in RTSys.

Medcom: a sustainable collaboration in Denmark

The Medcom National Health Care Network (Medcom) was established between 1994 and 1999, after several pilot projects that started in a regional scale (in the island of Funnen) by the late 1980's. In 1999, Medcom became a permanent organization joining the Danish government, local authorities and IT companies (Jensen *et al.*, 2004). This network is based on the exchange of standard clinical messages. By messaging each other, multiple systems from different vendors running on different institutions can exchange clinical information, from discharge letters to prescriptions and lab results. By 2010, near 4 million messages per month were exchanged and some of the tokens were close to 100% coverage (especially primary sector communications with the secondary care).

The MedCom collaboration, owned by the national health ministry and regional and local authorities, became one of the most relevant case studies in collaborative clinical messaging worldwide (Deutsch *et al.*, 2010).

Massachusetts eHealth Collaborative: private and public funds to structure collaboration

The Massachusetts eHealth Collaborative (MAeHC) initiative (MAeHC, 2012), formed in 2004, installed a RHIN providing health information exchange support to connect care in three different communities, covering over 500 physicians and 500,000 patients (Goroll *et al.*, 2009). As part of the strategy, MAeHC provided EHR systems free of change to the physicians, able to link clinical and administrative practice to form a coherent global solution. An explicit opt-in approach is used, in which patients are expected to express their consent to the 'as-needed' communication of clinical data between parties (reporting an impressive 90% of joining by patients).

The network is structured around three main services: the Record Locator Service (RLS), Clinical Data Exchange (CDX) and an ePrescription gateway (Halamka *et al.*, 2005). The RLS provides a community wide patient index, with no clinical information cached. However, links are available to the places in which patients have received care (a concept also used in our RTSys). The CDX provides the capabilities for electronic exchange of clinical information, including medication and allergies. The ePrescription Gateway supports the exchange of prescription data and coordinated workflow among applications from multiple vendors.

With the three pilots being rolled-out in 2005, there are some key success factors identified (Halamka *et al.*, 2005), including the option for 'coordinated decentralization' instead of a centralized approach. The proponents argue that local institutions are the proper stewards of the information, often reluctant to transfer the operation of their clinical databases to others. The data is kept inside each health care institution domain and the federation mechanism introduces the ability to aggregate it in controlled ways, a model that proved to deserve the agreement of patients, care providers and legal representatives (Halamka *et al.*, 2005). The option for a federation of decentralized data sources is also supported by the guidelines available from the

Connecting for Health think-tank, an initiative promoted by the Markle Foundation (Markle-Foundation, 2012).

Other initiatives in US and the stimulus pack

HIEx is on the political agenda in the US and much attention is being devoted to standards and interoperability. However, communities willing to implement bridges between competing systems still face hard challenges (Goroll *et al.*, 2009). Besides the case of the Massachussets, RHIN have been developed in many other states (Adler-Milstein *et al.*, 2011a); some pioneers include Indiana (McDonald *et al.*, 2005), New York (Kern *et al.*, 2007) and Maryland (2009). Lessons from the Santa Barbara RHIO show that finding a sound value proposition might be challenging (Miller *et al.*, 2007).

In 2009, the American Recovery and Reinvestment Act allowed Medicare & Medicaid Services to assign substantial financial incentives to those care organizations and professionals that use certified EHR systems in a 'meaningful way' (Blumenthal, 2009). The term (lacking a precise definition, deferred to latter regulation) implies the adoption of technologies enabling electronic prescriptions, electronic exchange of health information, and the ability to submit data for governance and clinical quality assessment.

Canada Infoway

Canada Health Infoway (Infoway) is an independent, non-profit organization whose members are the provincial and territorial governments. It was formed in 2001 by the Government of Canada to foster the use of EHR and information technology across the country, by promoting collaborative practices (Dennis, 2005). It provides a good example of how the interoperability in health information technology is a cornerstone to the national health system governance. Even if the goal is to define common standards and architectures for the entire

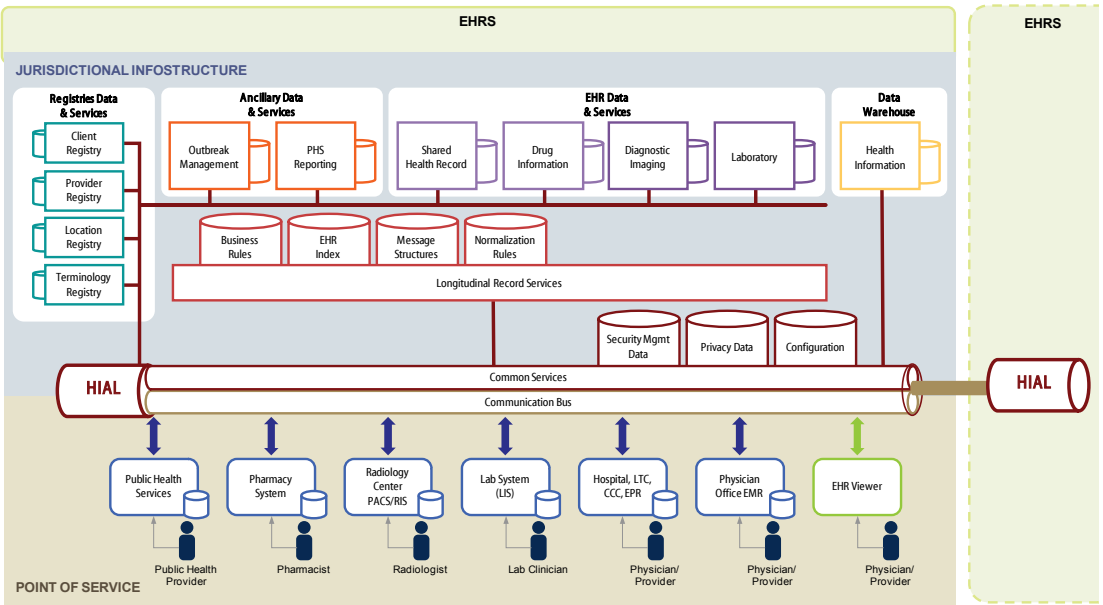


Figure 2.6: The Infoway EHR Solution architecture. Different EHR-S connect to each other to achieve national connectivity (adapted from (Infoway, 2006)).

country, the proposed models are useful to structure regional approaches, since the core problems are similar.

In 2003, Infoway presented the EHR Solution Blueprint, a framework that specified the required standards to support interoperable EHR systems countrywide, now updated to version 2, which expands the detail and scope of the initial architecture (Infoway, 2006). Every jurisdiction is expected to deploy its own EHR Solution, which must follow the enterprise architecture defined in the EHRS Blueprint, adhering to the same common information standards (Figure 2.6). Each discrete EHR Solution includes the standard EHR Infostruture and the heterogeneous Point of Service applications (these are the specific applications provided by different vendors).

The EHR Infostructure (EHRi) is external to the applications and provides the abstraction of a reference EHR system. All databases and applications should connect to the EHRi to participate in the information sharing space. Applications contribute with data to the EHRi and use it to access data publish by others.

Each jurisdiction maintains its EHR-Solution (which can be itself a network of smaller ERH-Solutions) following the Blueprint specification. Building the National EHR system corresponds to connect the regional EHR-Solutions, which are the authoritative sources of data, sharing a uniform model.

EHR Infostructure replicates data from Points of Service into four EHR specialized domain repositories (Shared health record, Drug Information, Diagnostic Imaging, Laboratory). All applications (bottom layer) maintain their own repositories; they are expected to talk through the HIAL (Health Information Access Layer), not to each other directly; they push copies of the local managed information to the EHR Data Repository.

Infoway specifications stress the need for standards to enable EHR systems interoperability; the 'blueprint' concept is itself an architectural standard providing a reference for the integration of new systems. International standards are adopted at multiple levels: messaging is based on HL7 v3 (HL7, 2012a); laboratory observations adopt (and extend) LOINC (LOINC, 2012); SNOMED-CT is used to code clinical concepts (SNOMED, 2012); IHE Integrations profiles (IHE, 2012) are used to specify how IT systems can engage in integration use cases.

Austria

As an example of another early adopter, the project *Health@net* is being deployed in the Tyrol region, in Austria, since 2002. This project implements a RHIN providing the electronic transmission of discharge summaries from the Innsbruck area hospitals to primary care providers in a unidirectional way, having supported the transmission of approximately 40.200 letters between June 2003 and October 2004 (corresponding to 8% of the total letters issued). A bi-directional possibility is planned through a web portal (Schabetsberger *et al.*, 2006).

Finland

A RHIN has been implemented in the county of Uusimaa (Finland), linking primary, secondary and tertiary care (Harno *et al.*, 2006). The proposed system enables shared EHR to

support ‘virtual workspaces’ of collaborating multi-professional, multi-site teams. The integration view is supported on a specific middleware, encompassing EHR location services (based on a link repository approach) and security and consent management. The information is kept at the original sources and explicit informed patient consent is required to access it using the middleware.

2.4 Portugal: the healthcare system context and e-Health initiatives

Portugal is seen as an average adopter of health information technology, when compared to other European countries (Dobrev *et al.*, 2008). Core capabilities, such as the use of computers in practice and network connectivity are available (UMIC, 2010). These capabilities, however, do not have expression in the continuity of care, and the country lags on the use of more advanced value-added services, such as electronic sharing of clinical data (Dobrev *et al.*, 2008) and on the definition of a comprehensive strategy for the use of health information technology (Ortega Egea *et al.*, 2010).

2.4.1 OVERVIEW OF HEALTH INFORMATION TECHNOLOGY USAGE

The Portuguese health system comprises both private and public care providers, but the public National Health Service (SNS) dominates the sector (Barros *et al.*, 2007). The SNS covers the majority of Primary and Secondary care facilities and is by far the major EHR keeper.

Health information technology in Portugal is profoundly fragmented, with a generalized lack of health information systems integration (Velez Lapão, 2007). This happens inside large organizations, in which departmental information systems are usually not interoperable (Cruz-Correia *et al.*, 2005), and across institutions. Up to date, no national production EHR system exist and electronic clinical communication for the continuity of care is limited to a few bottom-up projects (RSE-WG, 2009b).

Until recently, the public body in charge of national health IT governance was also a solution provider, accumulating with a software house profile. Solutions supplied from this body were provided at reduced cost to the public health institutions, dominating existing implementations. As a result, in Primary care the reference information system for patient management and basic EHR is the SAM-SINUS and, in Hospitals, the SAM-SONHO, although others exist. While these systems ensure the essential EHR functions, departmental solutions are also available, usually from other suppliers. In a large Portuguese hospital, the count of departmental information systems was estimated to exceed 60 (Cruz-Correia, 2008).

The private network of the Health Ministry, the Rede Informática da Saúde (RIS), assures connectivity between health care providers. RIS is isolated from general Internet traffic and uses leased lines from service providers. Broadband internet connections are available in 95% of the Hospitals (UMIC, 2010), allowing them do teleradiology (17% of the hospitals) and teleconsultation (11%) in routine.

Despite the good connectivity provided by the RIS, standard and routine electronic exchange of clinical information in the health system for the continuity of care is not implemented (Monteagudo *et al.*, 2007), especially with respect to the EHR content, which remains private to each health care provider (recall Figure 1.2, p. 3).

We believe Portugal should invest in information technology to achieve a more connected care, for increased efficiency of the health care system and quality of patient data available for clinical decision-making. This line of though, contrasting with the current practices, is also supported by other research groups in Portugal (Espanha, 2010; Cruz-Correia, 2008).

2.4.2 THE RTS PROJECT AND RELATED INITIATIVES

The RTS project was introduced in the context of the *AveiroDigital* program as a R&D project to develop new online services to connect the patients and health care providers (presented in section 1.2). In face of the incipient state of e-Health development in Portugal at the time, the RTS brought an innovative approach, facilitating computational methods for institutions to share health information, preserving all their responsibilities and autonomy. Other initiatives aiming at connected health care in Portugal were independently developed by different projects and briefly reviewed below.

The 2009 national health record initiative (RSE)

A large number of countries have implemented or are working towards some kind of national EHR system (materialized or virtual), which is a patient record contributed by multiple care providers, allowing for a comprehensive and lifelong clinical view of patient (Deutsch *et al.*, 2010; EC, 2007c). Such resource is not yet available in Portugal. The main reason for this is the absence of a regulation and technical framework for health information sharing.

Profound changes were expected from the *Registo de Saúde Electónico* (RSE) initiative, launched in 2009, which set up a task force for specifying a EHR interoperability framework in Portugal, much driven by the need to comply (as a Member State) with pan-European plans on eHealth (EC, 2004; European Union, 2011), with direct impact on EHR systems interoperability (EC, 2008a). The task force, however, was terminated in 2011, in the context of the financial crisis affecting Portugal, without having the opportunity to complete its work.

It is interesting to note that some foundational concepts in our work, initiated in the context of the RTS project in 2004, are also present in the planned architecture for the RSE system (RSE-WG, 2009a). The RSE specifications proposed an EHR interoperability framework based on a federated architecture, with a normalized and stable shared subset of the EHR. This somewhat confirms the work developed in the RTS project.

Regional initiatives for connected care

A few bottom-up initiatives exist in Portugal aiming at articulating the shared use of clinical data.

***Informação Clínica do Utente* (ICU).** The ICU was developed in Hospital S. João, a very large Portuguese Hospital (over 1300 beds), to provide a single point of access to clinical reports (ICU, 2012). The system adopts a Virtual Patient Record approach to create the abstraction of an integrate access to the patient data collected from eleven heterogeneous departmental systems (Cruz-Correia *et al.*, 2005). Even if inside the same institution, many of the integration problems occurring in RHIN can already be observed. The system is structured in three main components: the web-based visualization layer, the multi-agent integration system, and a central repository. The later holds the integrated clinical documents and metadata for version control. The documents are materialized in the integration view using HTML and PDF formats. In 2008, the system had 3 million documents in the repository.

***Processo Clínico Electrónico Único (Madeira)*.** The political status of Madeira allows it to have its own health system organization. This has facilitated the setup of a regional health information network. Madeira chose a centralized approach, with a scalable data centre providing a unique, digital EHR for the region, the *Processo Clínico Electrónico Único* (PCEU). The information available in PCEU is divided in administrative and clinical subsets, supporting the needs for Primary and Secondary care, both in emergency and planned care. The system provides role based authorization and allows different subsets of the information to be accessed, providing, for example, just a summary view to technical supporting staff. In order to evolve towards the progressive use of digital-only clinical records, a supplementary project worked on the digitalization of existing records.

***Urgência Pediátrica Integrada do Porto* (UPIP).** The regional administration of the Portuguese Health System redefined the referral scheme for paediatric care in the metropolitan area of Oporto, which motivated the development of a new supporting information system, with regional, multidisciplinary purpose. The new UPIP system provides a central repository with a subset of administrative and clinical data expected to be shared by all the institutions engaged in the paediatric care (*e.g.*: list of problems/diagnosis, chronic medication, prescribed diagnostic procedures, clinical notes, basic administrative information on existing episodes). The system should also allow to access external systems with valuable information, besides the common dataset in the regional repository and the articulation with the reference systems in primary and secondary care (SAM-SINUS, SAM-SONHO) was sought. Besides providing access to information, the system also supports electronic referrals, mapping the paediatric workflows in the region.

2.4.3 PORTUGUESE REGULATION CONTEXT FOR HEALTH INFORMATION SHARING

The information about an identified (or identifiable) person is legally labelled as personal data. The protection of personal data is a fundamental right, already recognized in the Convention 108 of the Council of Europe (CoE, 1981). In Portugal, it is regulated by Law 67/98 (1998), which transposes for the Portuguese legal context the European Directive 95/46/CE (European Union, 1995). This Directive provides the most important regulation in data protection for the present discussion and its purpose is two-fold: to ensure the proper protection of individuals' privacy and to enable a (regulated) flow of information in Europe, in support of

market development. The ruling aims, therefore, not at stopping electronic information flow, but providing a framework for its best use.

The Directive identifies the concept of 'special sensitive' data; examples include health care data, sexual life or genetic data. As a rule, the use of sensitive data (including health data) should be supported by the subject of care consent, which, to be valid, must be informed, explicit and specific.

Portugal does not have a dedicated legal framework for e-Health (Monteagudo *et al.*, 2007) and, consequently, for electronic communication of patient data. In this case, the general data protection act applies (Law 67/98), supervised by the Portuguese data privacy agency, *Comissão Nacional de Proteção de Dados* (CNPD).

Being a fundamental right, privacy must be balanced with other fundamental rights, namely the right to live. In this context, the Portuguese law authorizes the use of personal data in health care without the need to explicit ask the subject of care for consent, as long as these two premises are observed:

1) There is a need-to-access: the information is necessary to establish a diagnosis or conduct treatment or for the management of service units.
2) The context in which the information is used provides adequate data protection: care professionals are obliged to professional secrecy; the specific system and purpose are notified with CNPD; suitable system-level security measures are in place by the data controller.

A network environment aggravates the threats upon data privacy (van der Linden *et al.*, 2009) and raises the question whether the patient data should even be communicated to another health care organization. Also in this case, the Portuguese law does not request an explicit consent of the patient, given that the access is required for justifiable treatment continuity. Here, we should distinguish between different types of information in the health record. While allergies, alerts, medication, diagnosis and other essential information to deliver safe care must not be omitted, the personal notes that a physician records should only be transmitted if s/he chooses to.

Several studies have shown that the use of ICT in healthcare does raise legal concerns (Doosselaere *et al.*, 2008). In addition to the legal dimension, there are other potential threats connected to poor data protection. We should note that a breach on the confidence on data safeguarding practices can lead the patient to share less information with health professionals, which harms best treatment assessment and safety. For the health professionals, a similar scepticism may lead them to under-document in the health record. But the use of health information technology can also be an opportunity to enhance data protection practices which traditionally leave a lot to be desired in the Portuguese hospitals (CNPD, 2004). In a recent ruling, the Portuguese Council of Ethics in Life Sciences provides a synthesis of the current views and draws a sound foundation for the development of ICT systems supporting clinical information sharing (CNECV, 2011). The number 10 in this ruling is particularly relevance for this work, and explains that a professional can share health information with a second professional (using health

information technology) if that is relevant to treat the patient. The subject of care (the patient), however, has the right to know when and by whom his record has been accessed.

The Portuguese legal context also identifies the data processors (*e.g.* Hospitals) as trusted custodians of health data, with the obligation of preserving the health record systems within the organization control. This means that central repositories of EHR collide with the Portuguese data protection practices.

3 Connected region for the continuity of care: the RTS application concept

Evidence shows that deficient communication of clinical information between care providers may impact patient safety (Kripalani *et al.*, 2007; EC, 2007b) and process efficiency (Walker *et al.*, 2005; EC, 2007b). On the other hand, the use of RHINs can enhance the availability of information to support teamwork between disciplines and providers (Mäenpää *et al.*, 2009). The *Rede Telemática da Saúde* (RTS) electronic platform was a bottom up initiative to improve clinical information flow in the metropolitan region of Aveiro, challenging the exiting fragmentation of health information technology and the lack of digital clinical communication.

In this chapter, we elaborate on the requirements defined in the context of the RTS project and the use cases selected to be supported by the telematic platform (the RTSys system).

3.1 A model for connected care in the region of Aveiro

Primary and secondary care organizations are responsible to set up and operate their own internal clinical data management processes. Investments in health information technology are traditionally oriented to the internal needs and little effort is put in the collaboration with other health actors. The current status is explained by the lack of regulation and governance frameworks to facilitate (or compel) organizations to address collaboration through electronic information exchange. In this context, the RTS project proposed the use of a digital infrastructure to connect partners in the public health sector and allow them to exchange part of the patients' clinical record, relevant for the continuity of care.

The proposed platform introduces new collaboration practices and, being a seminal approach in the Portuguese context, had to be carefully designed to ensure it met the best privacy and ethical principles, while enabling organizations to open their data to the others and to the patient. The good examples from other countries (Schabetsberger *et al.*, 2006; Katehakis *et al.*, 2007; Mäenpää *et al.*, 2009), more advanced in the use health information technology, and the successful research on regional health information networks in Europe (Oates *et al.*, 2000; Saranummi *et al.*, 2005), provided an additional stimulus to bring those models into the national reality.

The requirements for the *Aveiro*'s digital platform were jointly defined by the care organizations (representing two Hospitals and the primary care sector) and the University of Aveiro, in the context of the RTS project.

3.1.1 REGIONAL COLLABORATION PLATFORM FOR HEALTH PROFESSIONALS

The RTS project was motivated by an obvious gap between the collaborative nature of the practices in the health care system and the lack of information and communication technologies to support the flow of clinical data. No structured electronic communication of the clinical context of the patients between care providers and disciplines was available at the time (nor is it yet available nationally).

In the scope of the RTS project, we have analyzed the referrals occurring in the region, between primary and secondary care organizations (Figure 3.1); these results confirm that business processes (the care workflows) go beyond the walls of each organization (Haux, 2006), with special focus on the interactions between Primary and Secondary care.

The situation was characterized by a limited and paper-based exchange of information. Clinical information from the attending GP to the Secondary care specialist was conveyed in (tangible) referral documents; the feedback on Hospital contacts to the GP was based on paper-based discharge letters. This basic collaborative scheme presented important limitations: (1) no previous context on the patient was available when s/he attended emergency care (except for the information locally kept at that institution); (2) feedback on hospital episodes was late communicated to the GP (and often lost); (3) there was a perceived unbalanced relationship between care providers, with Primary Care having the obligation to provide detailed clinical context in referrals, but not the other way around for follow-up. Other generic inefficiencies that usually motivate partners to seek a RHIN (Mäenpää *et al.*, 2012) could also be observed, *e.g.* the possibility for repeated and avoidable diagnosis procedures.

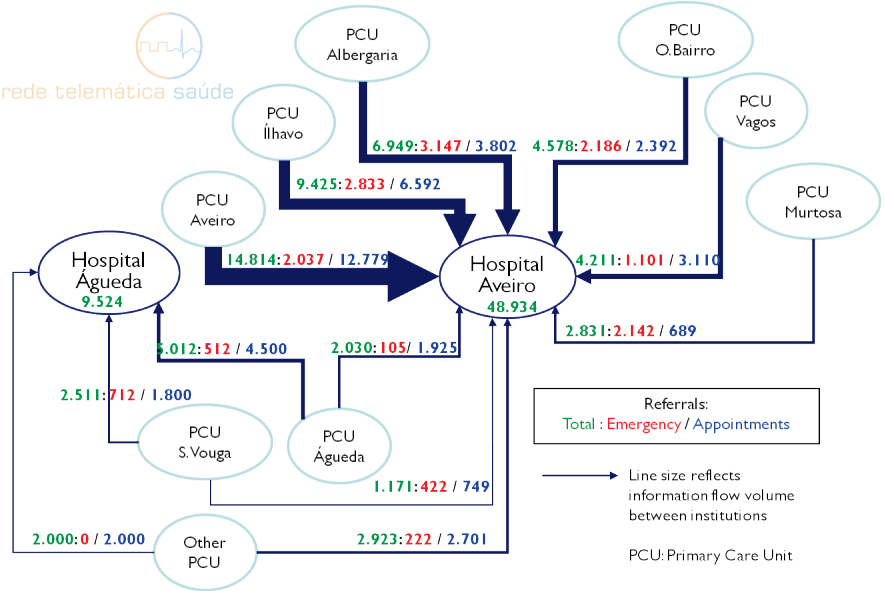


Figure 3.1: Flow of referrals in the metropolitan region of Aveiro (available from (Cruz et al., 2005)).

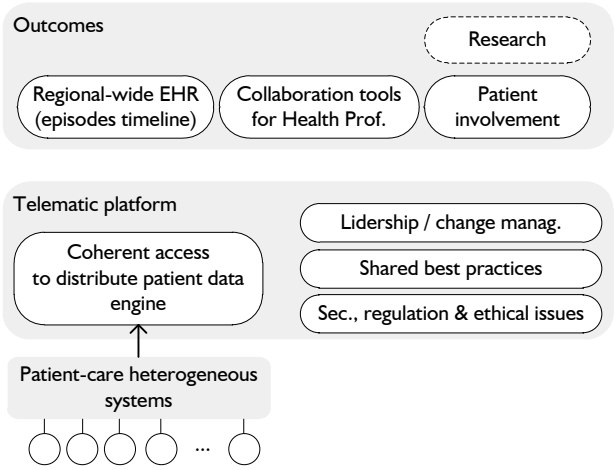


Figure 3.2: Overall vision of the RTS project: a telematic platform to connect health care practitioners to enable a regional wide EHR.

The awareness of the problems associated with limited information transfer between clinical institutions (Kripalani *et al.*, 2007; EC, 2007b; Kaelber *et al.*, 2007), and the lack of a national answer, motivated the project partners to look for computer methods that could enhance the circulation of pertinent information in the region, connecting the Primary and Secondary care.

The investment on this new RHIN was a bottom-up approach, in-line with the mission of the clinical institutions, but not originated in any Government directive. As such, there was a limited funding for the project activities and it was not realistic to expect that existing production information systems could be changed to adopt new models. The effort was to be directed towards the required bridging information and communication technology, enabling the shared use of clinical information already present in disparate systems. This concept of a new ICT layer, connecting different institutions, can be characterized as a telematic platform for collaboration, around which the all project was structured. The platform was expected to be used as: (1) a clinical tool for the health professionals, supporting the abstraction of a unified 'patient timeline'; (2) a collaboration tool to facilitate the communication (messaging) between professionals (Figure 3.2). The coordination of service points, explicit patient transfer procedures and all aspects dealing with resource allocation and costs were kept out of the scope of the project.

The first priority was then set at providing a practical way to access the timeline of the patient encounters in the region, with a clinical summary for each care episode. Additional services (*e.g.* a web channel for the patient to interact with the care providers in the region) were planned, with a lower degree of priority. A complementary outcome of the platform, as it was understood at the project inception phase, was the ability to use it for research, benefiting from the new information source (the regional aggregated view), composed by more complete patient cases (resulting from the information gathering).

A platform that opens the care information from one organization to another is not just an ICT challenge (Berg, 2004; RCGP-HIG, 2009). It must involve the strategic leadership of the management boards, the adoption of change management practices, agreeing on shared practices between the participating actors and the set up of mechanisms ensuring the trust of all

stakeholders (*e.g.* designing for security, privacy and legal issues). In this work, rather than defining all the multiple aspects supporting the setup and operation of a RHIN, we are focused on the ICT solution to the coherent access to distributed patient data (Figure 3.2).

3.1.2 THE ROLE OF THE PATIENT

The modern conceptions of the health care systems recognize the importance of putting the patient in the centre, meaning that the health care system organizes its resources and procedures to best serve the patient, and not the other way around, in which the patient is forced to adapt the health system idiosyncrasies. This line of thought is embodied in the expression patient-centred care (Philips-EIU, 2007; Shaller, 2007). There are several definitions available (Cronin, 2004), but we would like to retain the one from the National Health Council, proposing the idea of 'quality health-care achieved through a partnership between informed and respected patients and their families, and a coordinated health-care team' (Cronin, 2004). The patient should, therefore, be involved, through information, and have an active voice in the care process; healthcare teams should coordinate to serve the patient.

There are nowadays many Patient Health Record initiatives promoting the active participation of the patient. A Patient Health Record (PHR) is a record controlled by the patient, which he may configure (*e.g.* filtering access levels), structure and partially contribute (*e.g.* with self performed vital signs assessment) (Tang *et al.*, 2006a). To be effective and useful, the PHR should be comprehensive and collect, as seamlessly as possible, information from existing distributed systems (Detmer *et al.*, 2008).

PHRs are an active and important area of current research in medical informatics (Kaelber *et al.*, 2008) but not the object of the present work, although there are a few points of contact. PHR methods stress the need for the integrated access to information stored in remote care information systems (Detmer *et al.*, 2008). In this sense, RTSys methods already provide (1) a web-channel for the patient to interact with the care system, and (2) a system for aggregation of information from network resources presented seamlessly. The RTS perspective is rather focused on facilitating information to the care teams.

The patient can participate in the RTS platform in several ways: s/he can (1) browse a limited subset of the his 'timeline', as known by the RTS; (2) interact with the health care professionals and organizations, by issuing contact requests; (3) monitor the accesses made to his information in the platform. Currently, the insertion of clinical content by the citizen is not supported.

3.1.3 GUIDING PRINCIPLES

There are multiple approaches to structure a RHIN based more on strategic visions and values about the health system and the role of health information technology, than strictly on pure system development approaches. These fall into what we can call principles. The RTSys principles arise from the perceptions shared in our group (at the IEETA research unit from the University of Aveiro), building on several years of activity in the field (Figure 3.3):

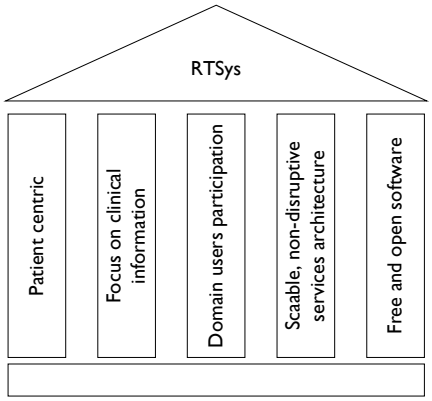


Figure 3.3: Conceptual view on the pillars of the RTS telematic approach.

P1: Patient-centric approach. Health information systems need to be coordinated and should facilitate an integrated view of the patient context to authorized actors, despite the service point in which the patient received care, to favour patient safety and best teamwork. The health domain must walk the path already taken in other industries and use ICT to support effective collaboration to serve the 'consumer' (the patient).

P2: Focus on clinical information. Administrative information has been the main concern of most IT investments in the past, but with limited results (Littlejohns *et al.*, 2003). The challenges are now around the best management of clinical information, to enable clinical decision, research and health policies (Aspden *et al.*, 2004).

P3: Specified with healthcare professionals. Given the complexity of healthcare (*e.g.* multi-professional hierarchic structures), ICT projects tend to be seen as problems and not part of the solution (Littlejohns *et al.*, 2003). A project such as RTS has to be strongly based on 'participatory design' (Dix *et al.*, 2004), *i.e.* on the participation of a representative number of healthcare professionals, so that complexity can be manageable and they feel that what is developed is also of their own. This approach has been previously adopted in our group, in the context of the Team-Hos project (http://www.ieeta.pt/team-hos/), which was selected as an example included in the text book by A. Dix (Dix *et al.*, 2004).

Participatory design raises the chances of success of the project as the system specification will be of higher quality, reflecting the needs at the point-of-care and, at the same time, users will acquire a sense of ownership and act as 'ambassadors' of the new telematic approach.

P4: Scalable IT architectural model that is inclusive and favours innovation. A RHIN system has to be backed up by a re-engineering process of the way IT solutions are provided. This presents an opportunity to define new organizational models for interconnecting practices, with the RHIO acting as a local 'regulator', ensuring the quality-labelling of IT solutions willing to participate in the sharing platform. With clear specifications and a system level architecture of pluggable services, the market would be able to contribute to the RHIN with compatible and innovative applications.

P5: Use free and open software. The complete RTSys infrastructure providing clinical collaboration services is implemented on free software, without the constraints of present and future licensing policies of proprietary software development platforms.

This set of principles provide a pragmatic orientation resulting from our own perspective and experience. A comprehensive discussion on principles to structure a shared EHR system, helping the reader to understand the domain and the consequence to the system design, is available from the British RCGP health informatics group (RCGP-HIG, 2009).

3.2 Specification methodology

ICT projects in health care often overlook the domain complexity (Littlejohns *et al.*, 2003; Dobrev *et al.*, 2009; Berg, 2001). A project such as RTS, which is expected to affect the teamwork and daily practices, should take the care to accommodate multiple points of view, respecting the diversity of disciplines, organizations and roles. Involving the end users in the specification teams, collaborating in the design of new work models (rather than being passive spectators), we raise the chances of success of the project and the quality of the analysis.

The specification activities in RTS were organized according to the guidelines of the 'Best Hospital Practice' Methodology, or BHP Methodology for short (Serrano *et al.*, 2000; Cruz *et al.*, 2002), developed in the European project Team-Hos, in which our research group has participated. The BHP Methodology is particularly focused on process redesign to enhance teamwork though effective ICT. It starts by assessing the health care organization against an abstract, optimal set of practices to detect gaps, in which special attention is devoted to teamwork activities, those originating collaborative workflows between different service points and partners (Serrano *et al.*, 2000). The next step is the process redesign, in which the definition of appropriate ICT support is an intrinsic part of the method.

The RTS project applied the BHP Methodology according to its three sequential phases: Analysis, Specification and Implementation. During the Analysis phase, the collaboration processes and information flows between different healthcare institutions in the project were studied, to identify the best candidates for improvement. The subset of selected processes were further detailed during the Specification phase, in meetings carried out by nine multidisciplinary teams, integrating healthcare professionals of different institutions, coached by the University researchers. These workgroups involved 42 professionals (17 doctors, 14 nurses and 11 administrative employees), over 60 working meetings (RTS, 2005). All teams received basic training on the specification techniques, specially on the BHP lifecycle and activity modelling using the UML notation (Booch *et al.*, 2005).

The work of the specification teams was organized in pilot areas according to the results from the Analysis phase. For each pilot, one or more teams were set up, representative of the relevant actors and disciplines. The teams were asked to develop specifications focused on the regional teamwork, identifying, from the point of view of the professionals in the field, practical bottlenecks to solve (Table 3.1).

Not all the new work models resulting from the specification activities were supported by high-level decision makers. All the referral processes, for example, were dropped in a later stage, because it was argued that the National Health System would provide a solution for it in the future. Other practical implementation issues have also contributed to limit the scope of the processes selected for improvement. For example, the access to external systems used in routine turned out to be much more complex than expected (especially with respect to the main patient management systems), and some information sets, such as the clinical alerts registered in Primary care, were not accessible.

The third phase of BHP Methodology —Implementation— provides high-level guidance on the workflow and expected outputs for the technology introduction step, but does not include

Table 3.1: Initial pilots of the RTS project and the corresponding specification teams. Shaded boxes are supported in the RTSys telematic platform.

Pilot areas	**Teams**	**Teamwork-based processes to be supported**
Pilot 2: Summary Regional Health Record	Episodes summaries: admission, discharge and transfer.	Professionals' Portal : access to admission summary (emergency care) Professionals' Portal : access to transfer summary (emergency care) Professionals' Portal : access to discharge letter and transfer summary (in-stays)
	Lab and radiology results	Professionals' Portal : access to imaging reports Professionals' Portal : access to lab analysis reports
	Vaccination chart	Professionals' Portal : access to vaccination chart Professionals' Portal : inform new inoculations (to primary care unit in change of the patient) Citizen's Portal: browse vaccination chart and upcoming events.
Pilot 3: Resources Booking (region) and Nursing	Appointments requests	Professionals' Portal: primary care centre creates appointment in the Hospital Professionals' Portal: Hospitals create appointment in primary care Citizen's Portal: patient requests an appoint with GP.
	Nursing care	Professionals' Portal: access to nursing discharge letter. Citizen's Portal: patient requests an appointment for nursing care
Pilot 4: Healthcare Professionals Collaboration	Collaborative messaging	Professionals' Portal: Secure exchange of documents. Professionals' Portal: Cases discussions and exchange of practice notes.
Pilot 5: Portal for the Citizen	myRTS: private area for patient requests	Citizen's Portal: browse "health agenda" and upcoming events. Citizen's Portal: browse position in waiting lists (e.g.: for future surgeries) Citizen's Portal: request GP for clinical advise Citizen's Portal: request GP to issue a standard report (e.g.: stereotyped medical certificate)
	Health education portal	Citizen's Portal / public area: publish health education content
Pilot 1: Infrastructure Telematic Services	Infrastructure	Professionals' Portal: identification of patients Professionals' Portal: authentication and authorization Professionals' Portal: audit trail Professionals' Portal: electronic authentication of documents Citizen's Portal: authentication and authorization Citizen's Portal: account and demographic data management

software lifecycle guidelines. For this purpose, the principles of agile software development (Larman, 2003) were applied, with the adoption of the Scrum method to structure the development effort (Schwaber *et al.*, 2001).

The active involvement of domain actors in the specification of the RTS functions allowed the promotion of shared best practices even before the implementation of the system. The best example is the consensus process towards a uniform discharge letter for in-stays that resulted from the 'episode summaries' team. For the professionals involved, this was a novel opportunity to discuss and compare practices with their peers at different institutions, towards better ways to collaborate and exchange information.

The specification phase confirmed the priority to build the Regional EHR since it provides the basis for other services that can be plugged in a later stage, as the platform matures.

3.3 RTS platform use cases

A RHIN is a collaboration space in which services with great latitude can be implemented (Mäenpää *et al.*, 2009). This means that an initial task was to identify the areas to be supported in the RTS telematic solution, within the period of the project and available resources, by establishing priorities compatible with the results obtained from the specification teams.

At the same time, the new ICT solution needed to be non-disruptive, in the sense that the existing information systems, liabilities and practices should be preserved. The opportunity for the RTS platform was to be found in the (1) cross-institutional processes, (2) focused on clinical needs, (3) within the range of the bottom-up approach adopted in the project. These orientations guided the joint specification teams to come up with a set of use cases for the health professionals (Figure 3.4), the first priority, and a set of use cases for the citizen, as an additional goal.

Use cases for the health professionals

Early in the project it was clear that if one was to enhance regional collaboration, then it was needed no enhance regional communication, promoting sharing practices, especially with respect to (1) sharing the clinical context of the patient, and (2) allowing professionals to use the platform as a clinical forum to help their practice.

With respect to the first objective, sharing the clinical context of the patient, the following use cases were selected (Figure 3.4):

— **UC1.1: Explore clinical history**. The authorized (health) Professional accesses the (RTSys) Portal, seeks for a patient and accesses the clinical history of the patient. This 'history' is a set of health characteristics relevant to establish the clinical context of the patient (and not documenting a specific encounter). The expected information would include alerts (*e.g.* known allergies), use of artefacts (*e.g.* glasses), laterality, etc. The system should retrieve information for any patient known to at least to one of the information systems being integrated.

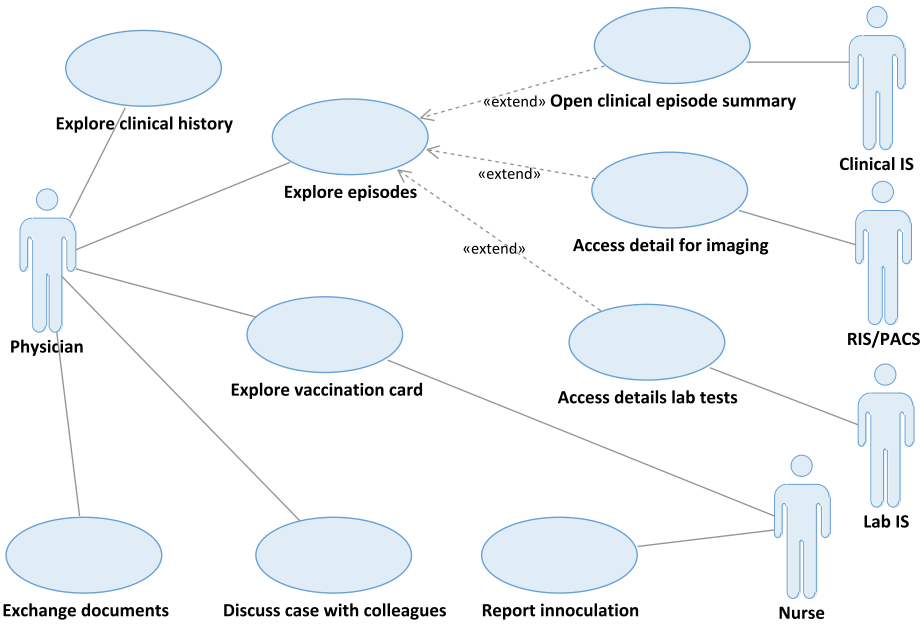


Figure 3.4: RTS platform use cases for the clinical professionals.

— **UC1.2: Explore episodes**. The health professional seeks for a patient and the system list all the known encounters recorded in the partner institutions. The list provides only basic information, reporting when and where the patient received treatment, and who was the professional in charge; in this way, the health professional can quickly visualize a timeline of the care encounters that occurred in the partner institutions. This view does not provide details on the health conditions or treatment received. The list should be organized according to different axes (*e.g.* events marked in a temporal line, episodes classified by type, etc.) and allow flexible sorting and filtering. This use case (Explore episodes) can trigger optional extension scenarios, depending on the type of episode being listed and the underlying details available.

— **UC1.3: Open clinical episode summary**. The Professional may choose to retrieve a clinical summary for a given encounter, formed by information extracted on demand from the source information system responsible for the documentation for that care episode. Though the information may come from different systems (*e.g.* episodes in two different Hospitals), the user should have a coherent visualization of the information (*e.g.* two discharge letters, coming from different organizations and systems, should keep the same fields and visual structure).

— **UC1.4: Access details for imaging** and **UC1.5: Access details for lab tests**. Some encounters originate rich data types, such as multimedia information. The access to these specialized information types may require high-bandwidth, specialized tools or specific skills. Given the particular constraints to visualize these types, they should be supported as extensions to the core scenario, and not required to form the regional EHR. This means that the information required to model, for example, medical imaging modalities or lab tests is not represented internally. The common denominator to these 'external' data

types is a text report, which should be available in RTSys. In some cases, the access to images and lab tests may be supported by redirecting the user session to external systems.

— **UC1.5: Explore vaccination chart**. The Professional browses the vaccination chart in the Portal, which displays the information registered for the patient at his primary care unit.

— **UC1.6: Report inoculation**. Some inoculations occur in care points other than the reference primary care unit for a patient. Travellers having a dedicate consultation in the Hospital, for example, may be inoculate there, according to the specific needs of the trip. Nursing stuff would them use the Portal to send a notification to the reference primary care unit, so the new information could be inserted in the proper information system. This electronic communication compensates for the fact that RTSys does not writes information in the source systems.

Other use cases were identified to promote a less structured discussion about clinical cases than the existing obligations with respect to the use of medical records. This would be a facility for online messaging, allowing direct linking to the patient context:

— **UC1.7: Discuss case with colleagues**. The discussion of cases is an electronic messaging facility, enhanced with the possibility to cite patient cases. This means that the target recipients would open the regional EHR of the patient of interest with a single click, inside the Portal environment.

— **UC1.8: Exchange documents**. A health professional may use the Portal to send electronic documents to another HP, listed in the directory of known users from the partner institutions. This is essentially a mailbox approach, accessible only in the secure environment of the Portal.

There are some simplifications with the use case diagram for enhanced readability (Figure 3.4). The associations modelled for the roles Physician and Nurse are the most common, but the appropriate level of access is in fact dependent on access control lists. This means that the Nurse may browse the clinical record, if the access policies are configured accordingly. Given the nature of the proposed system, almost all use cases involve distributed operations and the participation of external information systems; only the actors referring to on-demand access are shown. In addition, all scenarios require the users to authenticate themselves to start a session in the platform, which is not explicitly modelled in the diagram for the sake of simplicity.

Use cases for the patient

The ability to used aggregated information from multiple organizations and the potential to structure common practices among project partners, opens the possibility to offer new services to the patient in a friendly portal. The following scenarios include use cases oriented to information access and others to allow the patient to interact with the partner organizations (Figure 3.5). This web channel is the 'myRTS' portal.

— **UC2.1: Browse health agenda**. The patient, using a portal environment, browses an agenda (calendar) containing past and future events regarding episodes of care in the

partner institutions. The patient may get alerts and advice for preparing upcoming encounters.

— **UC2.2: Track position in clinical service queues**. As a transparency measure, the clinical institutions were willing to share with the patient details on the waiting queues for future clinical procedures, especially with respect to surgeries, which could take several months until occur.

— **UC2.3: Monitor health record access**. The accesses made to the regional EHR of the patient are available to the patient himself. S/He can then track which professionals accessed the record and take proactive actions, if appropriate. The professionals, on their side, know that the accesses can be inspected by the patient.

— **UC2.4: Request medical appointment**. The Patient submits a request for an appointment at a service point, especially for his assigned GP. The appointment is not immediately booked, as some brokering is required by human staff (and thus the word 'request'). A similar request could be issued for nursing care at the reference primary care unit (**Request nursing care**).

— **UC2.5: Ask for clinical advice/docs**. A messaging facility allows patients and his assigned GP to interact (asynchronously). Using the platform services the GP can grasp over the patient regional record, maybe to help with recalling the patient conditions. In addition to ask for advice, the patient would be able to request bureaucratic documents, such as medical declarations required for several trivial purposes (*e.g.* hunting permits, swimming pool attendance).

— **UC2.6: Inform vital parameters**. It was initially planned that the patient could report, on

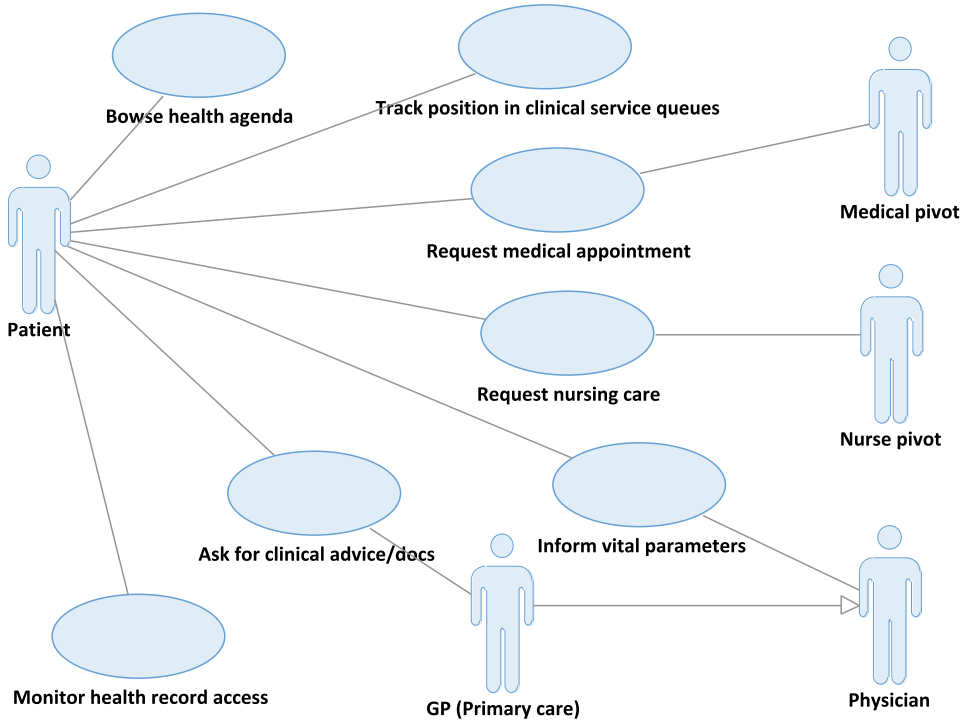


Figure 3.5: The RTSys use cases for the Patient ('myRTS').

his own, health parameters useful for the care relationship, including basic electronic Patient reported outcomes (*e.g.* weight).

Authentication is not shown in the diagram, for the simplicity sake. All use cases available in 'myRTS' require that the users authenticate in the system with individual credentials.

In addition to the myRTS services, the RTS project also identified the opportunity for a public health portal, with formative and informative purposes, edited by the health professionals from the partner organizations. Since these services should be supported by specific work procedures not related to the problem of serving aggregated information from the production systems (the technical challenge addressed in the present work), the public health portal will not be discussed here.

3.4 Requirements from the Portuguese context

Information technology for connected care is not trivial (Knaup *et al.*, 2007). There are challenging technical, organizational and regulation issues to be addressed when sharing clinical information in an interconnected infrastructure (RCGP-HIG, 2009; Stroetmann *et al.*, 2009). Several task forces have reflected on these problems and provided elaborated discussions on requirements and technical roadmaps, with different examples available from the United States (Markle-Foundation, 2012; Microsoft, 2009; Donnelly *et al.*, 2006)).

In the national context, the work by the RSE Working Group (RSE-WG) presents a conceptual solution for the federation of EHR systems in Portugal. The published report collects the contributions of several experts in health information technology with deep knowledge of the Portuguese health system (RSE-WG, 2009a). This work is posterior to the definition of the RTSys architecture, but departs from a very similar reality and draws compatible conclusions with respect to the requirements and general approach.

The set of requirements to build a regional interconnecting infrastructure is extensive and multi-layered. In the next topics, we highlight some of those requirements central to the solution concept and explain how we address them in RTSys, taking into account the specificities of the Portuguese health system and regulation environment.

Regulation environment requirements

In Portugal, the EHR should be kept at the guard of the responsible care institution and, therefore, stored and preserved by each organization. Electronic clinical data exchange between care service points is permitted but should be limited to scenarios of continuity of care (CNECV, 2011). Considering the existing regulation context and work with the Ethics committees in the context of the RTS project, we have identified the following practices to ensure proper data privacy protection:

— Demographic and Clinical data should be logically separated.

— The patient should have the means to verify the accesses made to his own record within the RTSys. The *a-priori* consent of the patient is not required for a professional to access an EHR, but every access is logged and can be monitored by the patient.

— The patient should be able to object to the sharing of his data between care organizations (opt-out), but the opt-in by default is coherent with existing practices.

— Clinical data is kept at the original sources (systems in which it was recorded) and no duplication should occur, except for the minimal data pointers that enable the localization of the remote data entries. RTSys global access is based on dematerialized views: when authorized professionals actually need the information, it must be retrieved from the original sources. Institutions keep the full ownership of the data they are custodians.

— Clinical data is presented only to professionals under professional secrecy code. Medical doctors (who already have access to the entire physical health record of all patients in paper) can access the shared record on RTSys, for all patients. It is assumed that a treatment relation with the patient exists, but such verification is not within the range of RTSys.

— The system enforces the use of Access Control Lists that define different access policies for configurable groups.

— All accesses are logged for further reference.

Data federation requirements

Data federation in RTSys involves the activation of distributed processes to obtain and relate clinical information. The following characteristics should be considered for the target reality:

— **Shared patient-identification by SNS number**. Patients share a common identification number, the national health system identifier that can be use to associate records in different organizations[3].

— **Read-only access**. Patient data is always entered in the specialized information systems already in operation, with proper security, maintenance and validation schemes. RTSys does not allow clinical data to be inserted or modified by users directly. Instead, the information must always be inserted in the source systems (keeping the existing practices) and the RTSys data discovery processes will pick the updates.

— **Autonomous processes**. The integration of the regional EHR should be supported by automatic computer methods, with online access to the source information systems, without requiring human intervention.

[3] A few patients do not have a SNS number. This is an unusual situation and they can be considered to be 'out of the system' (and not included in the RTSys processes).

System attributes

The telematic platform should meet non-functional requirements that ensure the extensibility of the solution, namely:

— **Build on service-oriented architectural styles**. The functions implemented by RTSys should be available as services, allowing the use of the infrastructure in future unforeseen scenarios that may benefit from the regional EHR to enable new applications (*e.g.* spatial-temporal analysis of patient mobility).

— **System level secure interactions**. All system level operations should apply a proper level of security, especially with respect to the communication between distributed modules.

— **Low impact on existing resources and the network**. The ICT infrastructure available in the partner institutions provides modest performance, with a poor network service level when considering the connectivity to the smaller primary care units. RTSys must not negatively affect the limited resources available.

4 RTSys computational model

The key outcome of the RTS Project is a regional health information network to support cross-institution care processes, enabled by a dedicated ICT platform (the RTSys solution). In this chapter, we discuss the RTSys design options, adopting the information and computation viewpoints provided by the Reference Model of Open Distributed Processing (ISO/IEC, 1998).

4.1 Overall architecture and key design options

There are several strategies available to overcome the lack of interoperability in health information technology and build the abstraction of a 'system of EHR systems' (refer to section 2.2). Each strategy corresponds to a high-level decision which constraints the system architecture options. In this sense, Table 4.1 provides a quick overview of the approach adopted in our work, by confronting alternative directions to design a sharing platform for clinical information.

The proposed RTSys design fall into what is often called an architectural approach to health information integration (Lopez *et al.*, 2009) in the sense that, instead of adopting a point-to-point clinical communication protocol, we define a new layer of services, abstracting the structure, extension and heterogeneity of the underlying information space. The new layer of integration enables a virtual regional continuum supported on three key abstractions: (1) a common

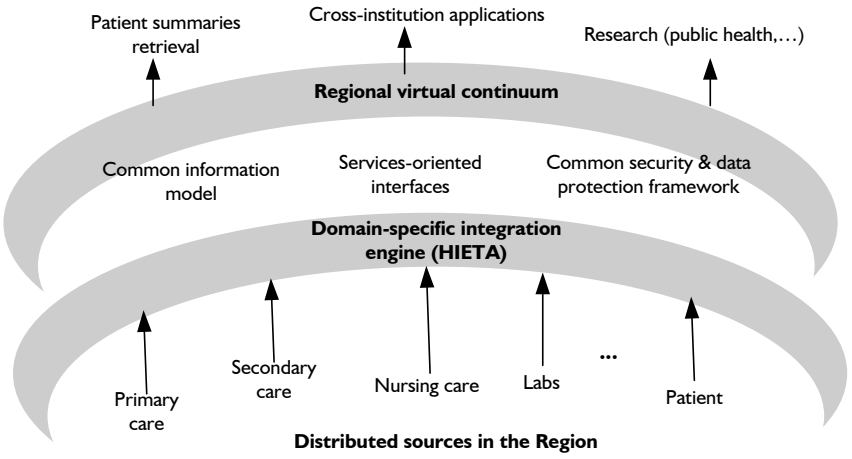


Figure 4.1: The HIETA middleware builds a coherent abstraction to enable regional applications.

Table 4.1: A summary of RTSys high level design strategies.

Architectural choice	RTSys design option
Information architecture:	
Central persistence vs. Distributed sources federation	RTSys is based on a federation approach, in which the content of remote sources is integrated to form a virtual, coherent view on the patient data. There is no persistence at the central node (except for lookup metadata) and the patients' information need to be retrieved from sources on-demand.
Patient summaries vs. Shared EHR	Only a part of the EHR is exposed to other partners. In this sense, RTSys adopts a summary view of the EHR.
Clinical documents vs. object models	RTSys defines and exposes an object model to represent the clinical data; it does not use formal clinical documents entities.
System architecture:	
Clinical message-driven vs. Services/components architecture	The modules of RTSys and external applications communicate via object-oriented services invocation. Data entities are represented according to the defined object model and marshalled according to the Web Services and J2EE specifications.
Presentation level access (web portals) vs. services level access (APIs)	Both are provided. The portals (one for the health professionals, one for the citizen) are examples of client applications built on top of the services API exposed in RTSys.
Standards-based interoperability profiles vs. Bottom-up approach	RTSys is based on a bottom-up approach, looking for the feasible convergence between the sources in use. The conformance to a domain interoperability framework was not set as a goal.
Domain specific vs. general purpose data integration engine	Processes dealing with extraction and transformation are specific for the health domain and the range of sources involved (though expansible to new sources). Domain concepts such as Patient and Episode play a central role to structure the distributed integration plans.

information model, (2) a uniform access through a services API and (3) a shared security framework (Figure 4.1). We use the term 'abstraction' since the continuum is apparent and arises from the action of a domain specific middleware, the Healthcare Integration Engine for Telematic Applications[4] (HIETA).

In this design, client applications are decoupled from the source systems and the way to explore the information resources is through the common, uniform layer, avoiding point-to-point integration scenarios. HIETA mediates between client applications and distributed data sources and provides a facilitator to manage the integration logic, providing well defined extension interfaces and ensuring the distributed operations. A relevant difference to the general deployment of Enterprise Service Bus architectures (Papazoglou *et al.*, 2007) is that (1) HIETA is domain specific, and (2) targets read-only scenarios (aggregates information from source systems).

The general architecture of RTS can be perceived as organized in three levels (Figure 4.1):

— **Level 1**: existing distributed information sources provide the required diversity to support daily practice. They implement specific semantics and access policies, and are deployed in heterogeneous IT. They differ on purpose and content, supporting departmental needs (*e.g.* radiology, labs) or the enterprise-wide functions (*e.g.* EHR systems).

[4] HIETA is a registered trademark of the University of Aveiro.

— **Level 2**: the RTS integration middleware provides the services to build a coherent 'regional view'. This includes two fundamental pillars of cross-institutional applications for health care: the abstraction of a common information model and a common security and data privacy protection framework. The integration layer exposes well-defined services to access and operate the regional view, such as listing the care episodes for a given patient available region-wide.

— **Level 3**: on top of the 'regional view' provided by HIETA, new applications can be developed, addressing the regional care provision value-chain. This includes, for example, sharing a virtual EHR, e-Referral schemes and implementing Order/Results circuits over electronic channels.

The key component in this architecture is the HIETA middleware (detailed in section 4.3.1). As the name reveals, the HIETA middleware is tailored for health care semantics on networked environments. It adopts an architectural approach, providing services and a uniform object model that abstracts applications from the underlying systems heterogeneity.

4.2 Information viewpoint

The RTSys adopts a common (or reference) object model to represent the target clinical domain and federates the distributed sources to feed this view. The federation principle implies layers of information (Sheth *et al.*, 1990) and in RTSys we distinguish between (1) the source information systems and (2) the virtual, region-wide health record. The regional record is expressed using the reference model, which is specific to RTSys and centred on the Patient, meaning that the entry point concept to the federated information space is the patient. Source systems are external to RTSys; they use heterogeneous information models along systems provided by different vendors, implemented using different technology. The remote information needs to be extracted and prepared to conform to the common model shared semantics, as implemented by the RTSys middleware.

4.2.1 THE VIRTUAL REGIONAL ELECTRONIC HEALTH RECORD

The key challenge with RTSys is to make distributed patient data available in a coherent, practical way, solving the three layers of interoperability: the system-level technical interfacing, the syntactic compatibility of data objects and the preservation of content semantics (Lenz *et al.*, 2007).

Interoperability issues are encapsulated by the integration engine (HIETA) to expose the abstraction of a virtual, regional Electronic Health Record (R-EHR). The 'virtual' concept is used to denote that users perceive a unified record but no materialized integration repository is kept at any central storage. 'Regional' here does not imply any special kind of jurisdiction or organizational arrangement: it is used to denote that the clinical context of the patient is contributed by several institutions, usually in a regional community.

Table 4.2: Information layers in RTS.

Layer	Repository	RTSys R-EHR scope
Minimal Data Set (MDS) layer	Cached in RTSys central catalogue.	Part I: Demographics Essential patient demographics (health system ID, name, birth date, gender, contacts, primary care GP and organization) Part II: Minimal clinical summary Clinical history (allergies, alerts) List of known contact cases region-wide (data, type, service point and health care professional in charge)
Common Information (CI) layer	Virtual (retrieved on demand)	Episodes/encounters: normalized data set for different care encounters (primary and secondary care). Medication available for each encounter. Vaccination chart.
Source information systems layer	Remote sources.	RTSys does not interfere with source information systems content or implementation.

Keeping the information at the sources (under the full control of its custodians) is a convenient approach to overcome data protection issues (Maro *et al.*, 2009), but it would be unpractical to discover all patient data on demand, as it would require always visiting the remote systems, exposing end-users to delays and remote systems availability, even for a simple patient lookup query. An indexing schema can alleviate this problem, working proactively on the discovery of patient data in the region to build a catalogue of the available information fragments and their location. The RTSys is able to fulfil each user request by first looking for the patient in the index and then, if required, retrieve on-demand the complementary data from relevant source system, where the clinical information is effectively kept. Similar approaches have been proposed in the literature, advocating the use of a fast-lookup index of the patient distributed data, tracking references to the source locations and not the clinical data itself, *e.g.* (Harno *et al.*, 2006; Tsiknakis *et al.*, 2002).

The RTSys 'index' (or catalogue) holds a minimal set of patient data that enables to identify the patients, enlist the episodes of contact in the region and retrieve a basic set of essential health characteristics of the patient. This information forms the Minimal Data Set (MDS), as defined by health professionals in the requirements engineering phase of RTS. In addition to this minimal view, which is in fact cached in the central catalogue, RTSys holds references (links) to the detailed information about episodes allowing, for example, to fetch lab results, imaging reports or even multimedia.

Patient information in RTSys is therefore multi-layer: the minimal data set is kept in a central repository and proactively updated by autonomous indexing processes; distributed fragments are kept at the sources and retrieved on-demand (but exposed according to a common information model); additional data is kept by remote systems, outside the scope of the RTSys processes (Table 4.2). The union of the MDS and the Common Information layers corresponds to the scope of the R-EHR accessible in the RTSys.

A central concept to the R-EHR is the episode of care and the use of summaries. The summary concept means that we are not aiming at feeding a full HER, rather access a subset

defined by a consensus process by the health professionals. The summary for an in-stay episode, for example, contains the discharge letter, diagnosis, prescribed drugs and the in-charge physician. This structure is the same despite the institution and information system providing the episode details. The discharge letter itself, being a complex fragment of the episode details, was subject to normalization within the project and a common layout was defined for RTSys by the health professionals.

The episode concept is already present in the source systems, especially in the leading Hospital information system (SAM-SONHO) with different granularities. For example, the in-stay episode can generate sub-episodes, which correspond to encounters at specific service points, such as radiology exams or lab results. The 'RTSys episode' is the coarse-grained encounter, aggregating sub-episodes to capture the care provided in multiple service-points, but related to the same initial contact (*e.g.* diagnostic procedures are sub-episodes of in-stay contact).

When a GP in the primary care unit accesses a Patient record, the central index is first queried to obtain a list of references of known episodes about the patient. The physician may then wish to open a summary of that episode, raising a request to the original information system, where the details are stored. A specialized adaptor module at that source would then extract and prepare the episode summary, as defined in the specification phase. The adaptor ensures the data transformation services required to bring the specific source content to the shared information model. The answer is then presented in the portal, in the same coherent interaction environment.

4.2.2 COMMON OBJECT MODEL

In the proposed data integration strategy, the content of remote sources needs to be adapted to conform to a common object model, following a Global-as-view approach (Florescu *et al.*, 1998). The common model is an object-oriented reference information structure, defining the domain concepts managed by RTSys. It is a reference model in the sense that it provides the stable representation of the domain, shared by all system modules (which is different from a very high-level reference model, providing the vocabulary to specify, by the use of constrains, the domain information structures, as in the ISO 13606 reference model).

The choice for a standard EHR content model (or for a standard framework to develop one) is a foundational decision to devise an integration platform. Adopting, for example, HL7 RIM (HL7) would imply the mapping of the existing content at remote systems to that model. Given that the systems integrated in the RTS do not offer HL7 compliant interfaces and their inner schemas were developed to address specific needs in different moments in time, it is unlikely that the mapping would be easy, if feasible at all (at least for some concepts in the model).

An alternative approach, as advocated by state of the art initiatives in EHR interoperability, is the use of two-level modelling (Garde *et al.*, 2007). In this case, the semantics of the EHR fragments can be defined using an archetype system (Wollersheim *et al.*, 2009). Since the archetypes (defining the structure of the information and the bindings of data members to controlled vocabularies) can be developed by any one working in the field, this approach is more flexible and can be tailored to the specific needs.

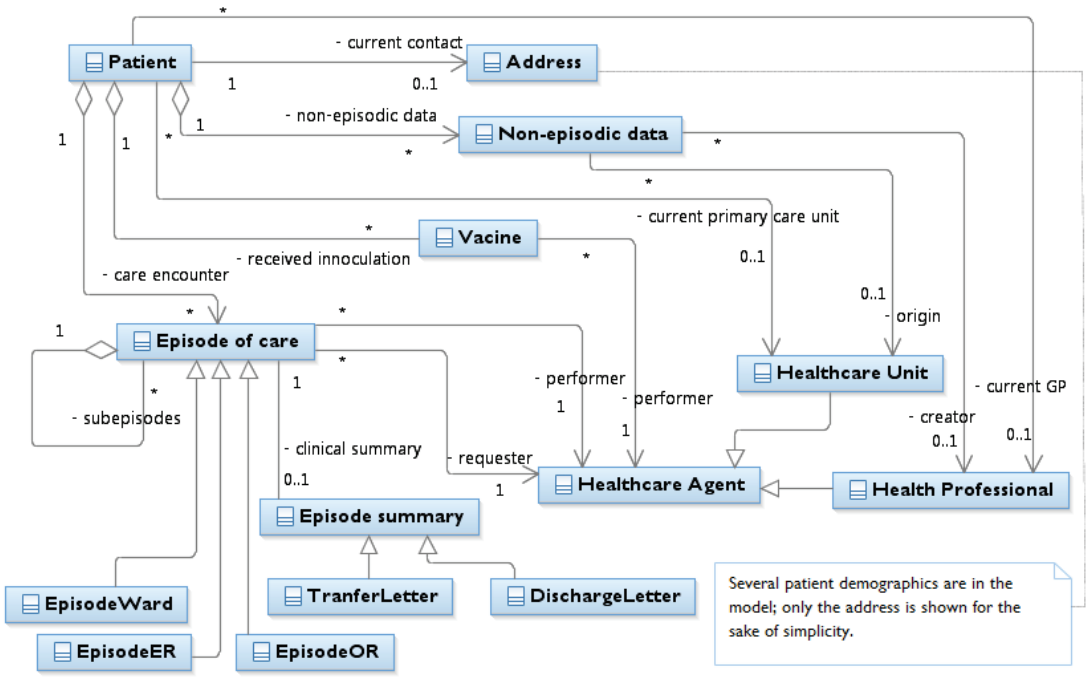


Figure 4.2: High-level view of the RTSys common object model.

In this context, the RTSys common object model neither is an information model used by an international standard in health information technology, nor adopts an archetype-oriented strategy. Instead, the decision for the RTSys content model was influenced by two main lines: the tangible reality of the information sources to be integrated, and our previous work in health information systems integration with the HISA common-components approach (Ferrara, 1998), in the scope of the HANSA European project. The European HISA standard, which has evolve to form the ISO 12967:2009 multipart standard, defines a reference domain information model, as discussed in part 2 of the standard (ISO-TC215, 2009b). The RTSys information model resulted from a bottom-up approach, departing from the available information sources and looking for the convergence with the concepts and relationships in HISA. The resulting RTSys information model is a viable object model for the content conciliation of existing systems, allowing us to have a practical implementation centred on supporting a shared summary record between institutions. It allows a (source) system independent representation of data entities, and enables an episode-oriented view of patient clinical data.

The RTSys common object model corresponds to the shared representation, at the federated layer. The central concept is the Patient (Figure 4.2). A Patient, which is the subject of care, holds identification, administrative and contact data (for the simplicity sake, the patient demographics are just represented by the Address concept). The Patient participates in care Episodes, an interaction between the patient and a health care institution, in the context of which s/he receives treatment. A certain episode may aggregate sub-episodes, for instance, an emergency contact can have associated laboratory and radiology sub-episodes. The episodes are classified according to the types in use in the SONHO/SINUS as Ward in-stay, Laboratory (lab work), Radiology, Emergency room care, Ambulatory, Scheduled appointment, Surgery (Operating Room) and

Medical Acts (a few types are depicted in the model, for illustrative purposes). Each Episode is described by an Episode Summary, which is materialized differently depending on the type of episode and the related domain practices (two sample subtypes are depicted in the illustration). The vaccination registry keeps track of the inoculation of Vaccines. A Healthcare Agent is a care provider, usually a Health Professional (namely, nurses and doctors), but sometimes used to represent a service unit. The Non-episodic Data represents specific clinical conditions of the patient, not pertaining to a specific contact (such as chronic conditions and allergies).

4.2.3 SOURCE INFORMATION SYSTEMS

The middleware approach adopted in RTSys implies that the information in the source systems should be extracted, aligned to the reference model and made available as data structures in response to services invocation. Since the external source systems are not controlled by RTSys, the ability to extract and relate heterogeneous information in a safe and sustained way depends on the availability of interfacing functions and documentation on those systems. If those interfacing extensions are not present, then only the collaboration of the source system providers can fill the gap, with clarification on the data semantics. In practice, this proved to be a limitation, since, in general, there were no programming interfaces available to interact with departmental systems (or they were not provided), and those interfaces to connect to the operational systems offered limited functionality. As an example, the most important information system in primary care (providing the patient demographics and GP encounters), the SAM-SINUS, exposes a very limited subset of its schema through documented database stored procedures (to collect scheduling information, vaccination and basic demographics). The system allows the GP to record clinical alerts and diagnosis, for example, but they are not exposed by any stable programming interface to external applications.

Information presented in RTSys proceeds from different systems (Table 4.3), and the same vendor solution may be instantiated in different institutions (autonomous instances). The information system *Sistema Integrado de Informação Hospitalar* (SONHO) is implemented in the two hospitals involved in the RTS project. SONHO is the most widely used hospital information system in Portugal, covering the generic functions of a patient management system. With the years, SONHO has evolved to support limited EHR functions and acts as a repository for two end-user oriented systems: the *Sistema de Apoio ao Médico* (SAM), for the physicians, and *Sistema de Apoio às Práticas de Enfermagem* (SAPE), for the nursing staff. The *Sistema de Informação para as Unidades de Saúde* (SINUS) provides the counterpart of SONHO, for the primary care centres. Similar to what happens in the Hospitals, there is a presentation layer tailored to the physician, the *Sistema de Apoio ao Médico* (SAM), working with the SINUS database (SAM-SINUS). The *Registo Nacional de Utentes* (RNU) is a national directory for patient identification. This system is somewhat recent and was not available from the beginning of the RTS project, and was inserted as a source in a later stage. The RNU is not always available (mainly due to connectivity issues) and its use, as for other sources, can be enabled or disable by configuration in RTSys. When the RNU is enabled, it is used as the most reliable source for patient demographics (name, address, GP and healthcare unit). IMAG is a radiology information system used in the HIP Hospital for certain types of medical imaging exams, developed in the University of Aveiro.

Table 4.3: Information sources integrated in RTSys (supporting structured information extraction).

Information system	Group of concepts	Concepts (detail)
SONHO (2 instances)	Patient demographics	Reliable source of patient identification for Hospitals. Patient has a local (record) number and a national identifier (SNS id).
	Professionals Identification	Identifies professionals with local employee numbers and the National professional association accession number.
	Episodes	Main source of patient episodes, which can be queried by patient and type. Provides associations to the time facts, professionals involved and summary (including diagnosis and discharge letter, if available). Episodes can be retrieved by type (Radiology, Lab, etc.)
	Clinical History	The data model allows to record information like allergies and family-related clinical conditions. The current practice is not to fill-in this information.
SINUS (6 instances)	Patient demographics	Besides patient demographics, one of the primary care centres can provide the GP in charge of a Patient. Patient has a local (record) number and a national identifier (SNS id).
	Professionals Identification	Identifies professionals with local employee numbers and the National professionals' board accession number.
	Episodes	The source for primary care episodes, both those that took place, and those planned (agenda).
	Vaccination chart	The vaccination record is obtained from the Patient's PCU.
	Clinical history	SINUS allows the GP to record some alerts, like allergies, that can be extracted to fill the clinical history.
RNU (Nation.)	Patient demographics	Patient Identity, contacts, assigned GP and PCU..
IMAG (1 inst.)	Radiology IS	Reports for a subset of the radiology exams.

There are three other departmental information systems accessed from the RTSys: Appolo and Clinidata XXI, two different laboratory information systems, and SiiMA, a radiology information system (deployed autonomously in two hospitals). For these systems it was not possible to have programmatic access to the information sources and the vendors only provided limited presentation layer integration.

4.2.4 SEMANTIC RESOURCES AND GAPS

The RTS information bus exposes a shared object model for applications. To this end, the local adapters need to align concepts and perform data transformations. Semantic mismatches may occur, which can be related to different causes, especially those related to the use of different terminologies and different data structures.

Terminology systems are available in the health domain to ensure data semantic preservation (Empirica, 2008) but their use assumes that they are sanctioned from an accredited domain body and, in some cases, that they are licensed. For example, vaccines can be coded according to the international SNOMED-CT and the 'Hepatitis A+B' vaccine would have the code

'333702001'. In Portugal, the SNOMED-CT is not licensed for the national health system and the practice is to code vaccines using the specific *Direção Geral de Saúde* controlled values (for Hepatitis A+B vaccine, it would be 'VHAB').

Semantic mismatches are more complex to solve when the concepts in the information models largely diverge. As the majority of source information systems integrated in RTSys were supplied by the public central administration of health IT, it was possible to take advantage on existing commonalities with respect to data sets and values sets to influence the design of the common information model and facilitate, in this way, the semantic concepts matching. All source systems, for example, share the national health system patient identifier, facilitating patient linkage.

In some cases, conciliation was not possible and different terminologies coexist, *e.g.* diagnostics in the SINUS and SONHO are coded using different terminologies, the first one uses ICPC-2 and the second ICD-9-CM.

The distributed data conversion occurs at the remote site. The logic required to extract the data from the underlying sources and align it according the common model is implemented as Java operations, *i.e.* methods in Java objects (Bloch, 2008). In a more complex scenario, with a large number of sources or with great variability of the data structures, adopting a strategy based on semantic mediation would bring greater flexibility. In this case, each participating source should be annotated, establishing the semantic associations with the reference ontology; the translation of queries (from the common representation into the site level) and merge of partial results could then be automated by using an appropriate semantic *reasoner*.

4.3 Computational viewpoint

From a system architecture point of view, RTSys provides a middleware solution to the integrating of health information from heterogeneous clinical systems deployed in the Portuguese health system. The middleware acts as a cross-enterprise information bus, allowing new sources to be plugged into the data integration processes, instead of the alternative point-to-point integration approach. The functional decomposition of the system, described in the next sections, relies on a federation approach, in which central modules plan and control the integration processes, and local adapters ensure the alignment of heterogeneous content into the shared semantics. Each remote system participating in the 'federation' is encapsulated by a specialized wrapper module, which exposes the underlying data as objects accessible through the invocation of Web Services (Alonso *et al.*, 2004).

4.3.1 AN INFORMATION HUB IN SUPPORT OF CROSS-INSTITUTIONAL PROCESSES

The central component in RTSys is the HIETA middleware that delivers a unification layer responsible for providing a common security framework and access policies, a common information model, and well-defined programming interfaces for applications development. Existing systems remain completely independent of the RTSys processes.

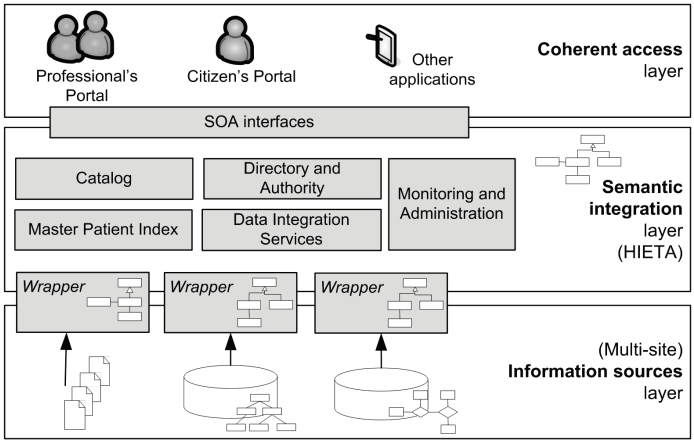


Figure 4.3: The HIETA middleware functional modules.

The HIETA middleware adopts a modular architecture, in the sense that the global functions result from the cooperation between specialized components, with well-defined purposes (Figure 4.3). The components provide a high-level decomposition of the solution, each embodying a reusable behaviour, relevant for the integration of networked information in the health domain. Similar component-oriented approaches to handle different aspects of health information systems integration have been advocated in the literature. We can find middleware components already in the CORBA extensions to the health domain, the CORBAmed (Blobel *et al.*, 1997); the so called 'common components' are at the heart of the European HISA Standard (Ferrara, 1998) and other health information systems interoperability initiatives, such as PICNIC/HygeiaNet (Katehakis *et al.*, 2007) and Synapses/Synex (Grimson *et al.*, 2002). In more recent approaches, we can find service-oriented architectures proposing the use of services that provide the functional scope of those long indentified modules; for example, HISA is now available as ISO 12967, adopting the SOA approach (ISO-TC215, 2009a).

The central layer in Figure 4.3 represents the components of the HIETA middleware implementing the integration logic, described below:

Wrapper modules work at each participating information source in the care network. They perform the semantic alignment of source contents with the shared RTSys reference information model, and expose the underlying data through Web Services. Wrappers are solicited both at integration time (for data discovery) and to supply on-demand user requests. They are also responsible for building standard episode summaries out of the underlying contents. The Wrapper abstracts the specificity of the source it was prepared to work with, meaning that, for a new source system, a new Wrapper should be prepared.

The HIETA **Catalogue** module manages the RTSys central index of discovered data. The Catalogue knows which clinical information is available and which is the (remote) system that manages it. In addition to a registry of remote data fragments, the Catalogue also keeps the required metadata to track the consistency of the information in the index and detect updates.

The **Master Patient Index** (MPIx) works in conjunction with the Catalogue to provide a reliable source of patient identification in the region. It adds domain logic to detect

inconsistencies in patient demographics and perform basic data cleaning. As an example, the MPIx distinguishes information sources quality with respect to patient identification, favouring those known to be more comprehensive and with more demanding data filling standards.

Directory and authority (D&A) services provide a reliable source of identification with respect to clinical professionals in partner institutions. This is a complementary function to the MPIx: while the later scrutinizes patients, D&A enlists professionals and service points. In addition, it also ensures authentication, role-based authorization and detailed user actions auditing.

Data Integration Services is the core coordinator module, triggering the distributed integration plans proactively. It is responsible for coordinating the dialog with Wrappers, passing them localized data queries and process the result sets.

The **Monitoring and administration** module provides a supervising toolkit, including user management and service level quality assessment. These services can detect variations in systems availability affecting RTS and adapt dynamically the behaviour of integration processes (detailed in (Santos *et al.*, 2009)).

HIETA's **Services interfaces** expose a stable set of functions available to applications working on the global data model. This enforces the access to shared data and functions through standard and documented middleware.

The bottom layer (Figure 4.3) represents the distributed sources. Each wrapper is expected to be deployed 'close' to the source it abstracts, and thus the Wrappers are active at remote sites. The top layer (Figure 4.3) represents the applications which use the HIETA to implement region-wide use cases. In the deployment of the RTS project, two portals populate this layer, one for the health professionals and another tailored to the patients. The logical layers depicted in Figure 4.3 can be mapped to different tiers in the solution deployment.

4.3.2 SERVICES FOR APPLICATIONS PROGRAMMING

A telematic platform that enables the global use of the patient data across organizations is a valuable asset for clinical applications (Mäenpää *et al.*, 2009). There are immediate use cases we can identify (*e.g.* serve a region wide EHR), but it also provides a foundation for innovation. This is a sound argument to expose the regional capabilities of RTSys using a services-oriented approach, enabling other applications to plug-in.

The services can be used to supply multiple applications, some under our control and others developed independently. As an example, in a complementary work developed in our group, RTSys was used to retrieve discharge letters in the region to feed an information extraction system, analysing the text of discharge records (Ferreira, 2011).

In a services-oriented approach, the services need to ensure:

— Well defined semantics concerning the services behaviour and the exchanged data structures.

Table 4.4: Web Services exposed for external application development.

Service	Purpose
Patient Index	Find patients (multiple criteria). Retrieve patient demographic data.
Health record	Retrieve the patient medical history (non-episodic clinical data) Retrieve the patient episodes and episode summaries.
Directory and Authority	Find health care organizations, service points and healthcare professionals. Retrieve and edit user accounts information.
Extended	Expose the information under other models than the HIETA reference object model (e.g. the system can export a patient summary in the epSOS format).
System diagnosis	Retrieve system status and statistics.

— Independence from client applications. The services are used in current use cases, but may also be in future applications, enabling innovative uses of the platform.

— Remote and secure invocation through Internet-friendly technologies and standards. The use of Web Services technologies is an established standard to enable service-oriented deployments, with the required security extensions to ensure protected communications and sound clients (consumers) authentication.

These requirements are met in RTSys but our architecture misses some more advanced features present in a true Service-Oriented Architecture (SOA), such as dynamic services orchestration and the use of service discovery facilities (Papazoglou *et al.*, 2007).

The RTSys services, implemented as secure Web Services, enable searching for patients, retrieve a summary regional record, authenticate and authorize users, and monitor accesses and system availability (Table 4.4). We can therefore think of RTSys services as programming interfaces to access a region-wide virtual repository of patient demographics, patient health data, list of service points and health professionals, authorizations and auditing trails.

RTSys services give access to personal information which must be accesses under proper policy enforcement mechanisms and the client applications are required to authenticate with HIETA and communicate over secure channels (Figure 4.4). The operations executed through the invocation of services follow the same security framework of those performed via portal (plus the need for the client application to be accredited): the user must login using individual credentials, which gives him access to different functions and information according to his role, originating an audit trail for eventual inspection.

4.3.3 COHERENT PRESENTATION LAYER

The use of a cross-enterprise information hub such as HIETA facilitates a single, virtual information repository for applications. This means that 'client' applications, such as the web portals developed in RTSys, interact with a uniform information model, despite the source system contributing with the particular data segment; an episode, for example, has the same data elements, independently of its origin. This enables coherent user interactions when exploring the

remote data, since the user is not exposed or required to learn different data structures or graphical interfaces, as they exist in the remote systems. The homogeneity in the content structure is also a convenient abstraction to support some more advanced user interactions, such as the use of aggregated data. An example of this is available in the RTSys portals, in which a flexible timeline summary view allows the user to 'zoom' events along time, which would not be possible if the web application would be exposed to the underlying schemas heterogeneity.

The use of HIETA as an information hub is a key to offer a user interfacing environment that displays the distributed information coherently. Another important option is the use of web technology. Basing the presentation layer in a web portal provides certain advantages relevant for the deployment of a regional solution (*e.g.* RTSys web applications are accessible from an ordinary browser, requiring no additional configuration). This is facilitated by the use of web development standards, as defined by W3C (W3C, 2012). The web environment supports the use of friendly rich user interfaces for appealing presentation, and provides a flexible approach to implement applications that adapt (changing the features available according to the user profile). In RTSys, there are different feature sets for the medical doctors, nurses, technicians and the system supervisor.

The option for an installation-free solution is even more pressing with respect to the citizen-oriented user interfaces, and the web technologies deliver the appropriate mechanisms. In addition, modern browsers support the authentication of users within web applications using *smartcards* (Zuquete *et al.*, 2008), which is required for the envisaged large-scale user login mechanism, based on the Portuguese national identification card.

The proposed architecture provides a 'clean' approach to explore clinical data, but will not work with remote systems that do not allow the access to the underlying data structures. These systems cannot be used as sources for the information hub, but still hold relevant information to complete the episodes details. This raises the opportunity to complement the basic RTSys presentation model —a web portal to access the federated information — with presentation level

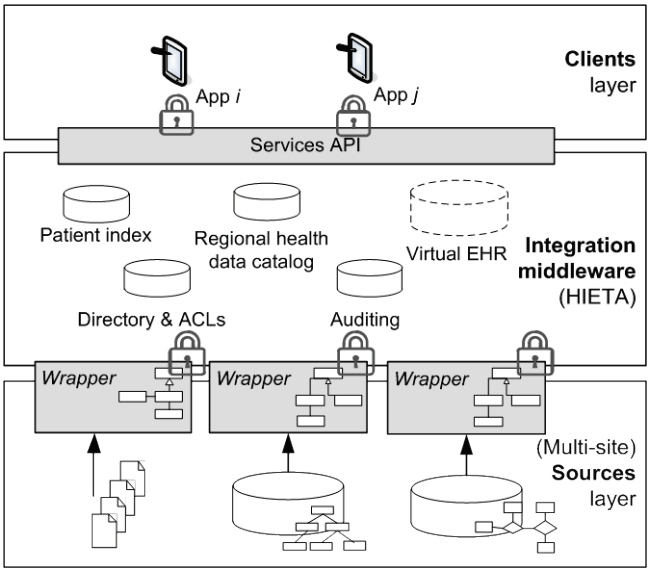


Figure 4.4: The middleware exposes a secure Web Services interface to client applications. The central modules interact with the wrappers over secure channels.

Table 4.5: Different access strategies to external information sources.

Information	Contributing sources	Access mode
Patient identification	RNU, SONHO and SINUS	Structured information extraction.
Medical history (alerts, chronic conditions)	SINUS, SONHO	Structured information extraction.
Vaccination chart	SINUS, SONHO.	Structured information extraction.
Episodes list	SINUS, SONHO.	Structured information extraction.
Episode summary	SONHO / SINUS	Structured information extraction.
Lab report	Appolo	Web access to external system.
Lab report	Clinidata XXI	Web access to external system.
Radiology reports / images	SiiMA	Web access to external system.
Radiology report	IMAG	Structured information extraction.

redirection. This requirement arises from the need to access external systems that (1) do not provide access to the underlying data sources but (2) are accessible through a web environment in the network. We can distinguish, therefore, between two different strategies to access remote content (Table 4.5): the RTSys principal approach based on the extraction of interpreted information, and a compromise strategy, based on presentation layer integration. Note that in the later case, RTSys does not transport the information structures, just redirects the user to a web interface provided by the vendor. This should be understood as a remedy to facilitate the access to remote systems since, by design, it does not comply with the cross-enterprise data bus approach central to RTSys and it is not possible to ensure a seamless user experience.

The presentation layer of RTSys will therefore support a hybrid model. The principal approach is to use the information structures provided by the middleware and, when not possible, to allow the user with a session initiated to be redirected to an external site, if available.

The common denominator to the RTSys use cases is the patient, in the sense that information revolves around a selected patient. The user experience, in the presentation layer, should then be structured around patient contexts and preserve the 'selected patient' across applications, as much as possible. To enhance the user experience, it should be possible that when a patient context is selected in RTSys, the redirection to an external system would place the user in the same patient record. The same holds when RTSys is invoked from external systems; if the physician is browsing a given patient in SAM-SINUS, for example, then s/he should be able to link to RTSys and get automatically the same patient. The continuity of the patient context between applications is a basic requirement for the professionals' acceptance and has been implemented in RTSys, as discussed in section 5.1.1.

4.4 Security e privacy architecture

Clinical applications impose demanding requirements on data protection since they store and process private, sensitive data (CNECV, 2011). A network approach may aggravate data

protection concerns as it proposes the blurring of institutional 'walls' and requires rethinking the way data is shared and controlled (van der Linden *et al.*, 2009; Bergmann *et al.*). The goal of networked health information systems is to deliver the best patient data protection while preserving the benefits of clinical information sharing (van der Linden *et al.*, 2009).

The boundaries to health data treatment and sharing in RTSys have been designed taking into consideration the regulation context and the active participation of the clinical organizations representatives (see section 3.4). At the heart of the adopted data protection strategy is the fact that the patient information is kept at the sources and fully controlled by the custodian organization. The key options for a secure and protective environment adopted in RTSys are systematized in Table 4.6. RTSys is registered with the Portuguese Data Protection Agency (CNPD), as requested by law.

Patient consent in RTSys

The setup of cross-institutional health data aggregation raises concerns regarding the position of the patients (Simon *et al.*, 2009; Hoerbst *et al.*, 2010b). Some patients may have the expectation that the clinical data generated at an institution is not shared with others and oppose to the shared access, despite the advantages for the clinical processes (Bergmann *et al.*).

In RTSys, it is implicitly assumed that the patient who uses the health care system consents that the clinical teams engaged in his treatment may access his clinical record, despite the service points in which it has been generated. While it is not possible to establish with full certainty that a care relationship exists between a specific professional and a specific patient in a given moment in time, using only the information in the source systems, the health professionals using RTSys

Table 4.6: Overview of security and privacy design options .

Practice	**RTSys design**
Data processor	Data is kept at the source systems and existing Data Processors keep all previous liabilities. The processing of the resulting aggregated views falls into the responsibility of the consortium.
Patient consent	A general opt-in principle is tacitly accepted, aligned with the current practices in health data processing in the Portuguese health care system. Patients can explicitly opt-out.
Patient control of his data	The patient is informed of all accesses made to his record in the platform, but is not required to authorize them *a priori*.
User profiles that can access patient data	Different ACLs are support for different user profiles. Only clinical staff obliged to professional secrecy is authorized to access the regional record.
Treatment relationship enforcement	The terms of use of the platform require that the health professional commit with the use of patient data for justified treatment relationship. However, the system is not able to verify or enforce it automatically.
Audit trail	All accesses in RTSys originate an entry in the audit trail and can be tracked (who accessed what and when).
Network-level security	Secure protocols are used to access the web portals and remote wrappers. Services invocation requires secure protocols.

are required to agreed with the terms of use, which state explicitly that patient data access requires a treatment relationship. The system does not prevent eventual misuses of data by an authenticated Professional, but all accesses are documented to enable further auditing. This assumption mimics the clinical practice and was supported by the representatives of the clinical partners in the RTS project, and also presented to the CNPD.

In addition, RTSys supports the use of opt-out lists, allowing to completely exclude patients from having their data used in the platform. For these cases, there is no emergency procedure that can override this setting and put them again in the platform for a very specific purpose, often known as 'break the glass' scenarios (Ferreira *et al.*, 2006).

Using the portal available to the patients, they can review when and by whom the R-EHR was accessed in RTSys, and s/he can ultimately take an active role on motoring how the personal data is being used.

Strong authentication of health professionals

In a complementary line of research, an architecture for the strong authentication of health professionals in RTSys has been proposed by Zúquete *et al* (Zuquete *et al.*, 2008), in close cooperation with the present work. This architecture proposes a sound authentication and authorization mechanism based on a particular design:

1) The Professional uses a smartcard (storing his credentials) for strong authentication.

2) The Professional identification and roles (for authentication and authorization of operations with the RTSys portal) are embedded in a short-lived X.509 public key certificate (Housley *et al.*, 2002), issued by the Professional hosting care-organization. The use of short-lived certificates needs no Certificate Revoking Lists.

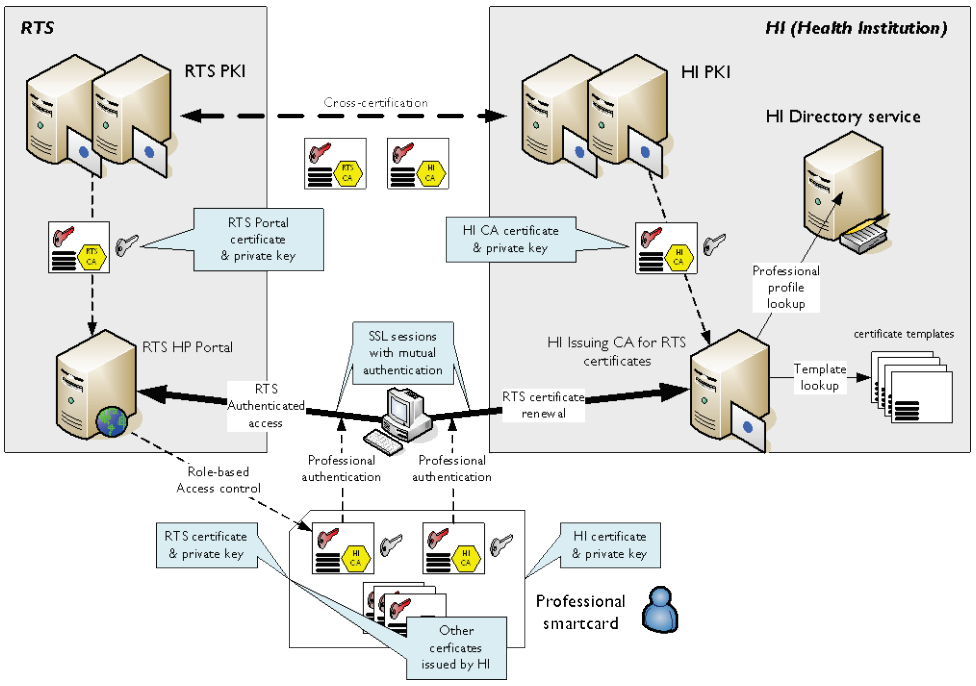


Figure 4.5: Overview of the strong authentication components for the health professionals (available from (Zuquete et al., 2008)).

3) The health care organizations identify themselves using "normal"-lived certificates.

4) All domains (the RTSys and each healthcare institution) run a private Public-key Infrastructure which is complemented with cross-certification to establish a 'circle of trust' among them (Figure 4.5).

The proposed design holds relevant features for the health care domain and regional partnerships: (1) the web technology (used in the RTSys portals) is ready to authenticate users relying on the client-side (browser) to get the user credentials in the smartcard; (2) each care organization operates a private Public-key Infrastructure and fully retails its ability and responsibilities to manage its users; (3) the trust relationships among partners is supported in cross-certification agreements and thus easily reconfigurable

Practical issues in securing RTSys

The use of the strong authentication mechanism described requires additional investments to which the partner institutions may fail to commit. As a fall back plan, RTSys allows the use of username and password authentication scheme too (which, ultimately, was the feasible option in the RTS project deployment).

The access to the R-EHR is only available within the closed private network provided by the Health Ministry (meaning that the professional is inside the 'walls' of a primary or secondary public care facility). This provides an additional level of trust when considering the protection against unauthorized accesses eventually through the Internet.

4.5 Abstract deployment models

The RTSys main target scenario is that of a RHIN in which partner institutions, connected through a private network, are willing to share information maintained in existing heterogeneous systems. In this mode, the RTSys acts as an aggregation hub, providing a region-wide layer for shared access. This was the actual scenario motivating the development of RTSys (in the scope of the RTS project), but other deployment scenarios can be addressed by the methods available in RTSys, such as (1) the federation of RTSys platforms, (2) management of shared care plans, (3) enterprise integration, (4) support to the regional service providers value-chain, and (5) a gateway to external interoperability networks. These scenarios are presented below.

The **federation** of RTSys platforms is a possible deployment in which some of the source information systems would themselves be instances of the RTSys (Figure 4.6). In this case, some sources would be RTSys instances, participating in a higher-level RTSys deployment. The semantic problems related to data alignment at the wrapper would be trivial to solve, since the schemas at the sources and the common model would be the same. This specific deployment has not been implemented but is compatible by design with the RTSys methods. A use case for such deployment would be the setup of a more comprehensive shared EHR, abridging Aveiro and other metropolitan communities, matching, for example, the administrative jurisdiction of the

Portuguese health care regions (Aveiro is a part of the Central Portugal administrative region), or even towards a national view.

The current design of RTSys does not support **shared care plans** in the sense that functions that allow defining, storing and tracking a care plan for a given Patient across multiple service points in the region are not supported. Users of the RTSys are 'limited' to a read-only access to a subset of the EHR. A possible evolution of RTSys could allow for central data repositories, besides the Catalogue, to support specific region-wide functions. The definition of care plans could them be supported in the RTSys Health Professionals Portal and saved in a new application repository inside RTSys. Though a possible and natural evolution of the platform, it would be a different use case, not centred on integrating distributed information as supported by the HIETA middleware.

RTSys can be used as an **enterprise integration** tool even if deployed inside a single organization, as long as there is the need to aggregate information from heterogeneous health information systems. RTSys could provide a simple approach to make information otherwise buried in legacy systems available in a friendly web environment.

A compelling reason to pursue RHINs is to optimize **the value-chain** connecting health care organizations (the consumers) and clinical service providers, implementing e-business methods between, for example, hospitals and medical imaging centres or blood analysis centres. The current implementation of RTSys is focused on the use of the HIETA middleware for information integration, and does not support the configuration and execution of cooperative workflows, *e.g.* forwarding the CPOEs to the providers and notify upon the availability of results. While this scenario is not fully support with current system implementation, a partial use case was adopted by the largest Hospital in the Consortium, HIP, with respect to the blood tests results access. In this case, patients are referred to the HIP for blood test by Primary care centres, using the normal channels (external to RTSys); the lab results are then made available in the RTSys, as a segment of the R-EHR (describing a lab episode).

A last example of a possible deployment scenario of the RTSys is its use to act as a single point of contact over a region to provide gateway services in a **broader interoperability scenario**. RTSys can be used, for example, to export the community-wide aggregated information to another integration layer/system, such as a National EHR initiative or a cross-borders initiative. This approach has been partially tested in RTS implementing the required services to act as a gateway to the epSOS interoperability initiative (epSOS, 2011).

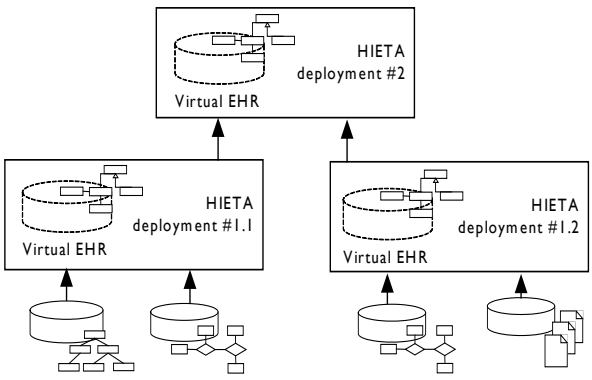


Figure 4.6: Schematic components in a hierarchical deployment of RTSys.

5 System implementation

The methods described in the previous chapter have been implemented using the J2EE standards (Alur *et al.*, 2003) and related technologies. In the present chapter, we discuss the implementation details, object interactions and technology choices. This chapter includes several screenshots to illustrate the implemented solution; please note that the data about patients and professionals in such illustrations is fictional and obtained from a development deployment, not connected in any way to the production clinical information systems.

5.1 RTSys production environment

RTSys was deployed in the context of the RTS project at the metropolitan region of Aveiro, connecting the two major Hospitals (HIP and HDA) and six Primary care units participating in the RTS project consortium (Figure 5.1). All the institutions belong to the Portuguese public health system and are connected to the RIS, the dedicated network for health care, managed by the Health Ministry.

The 'Aveiro deployment', known just as RTS (*Rede Telemática da Saúde*) is a first pilot of the

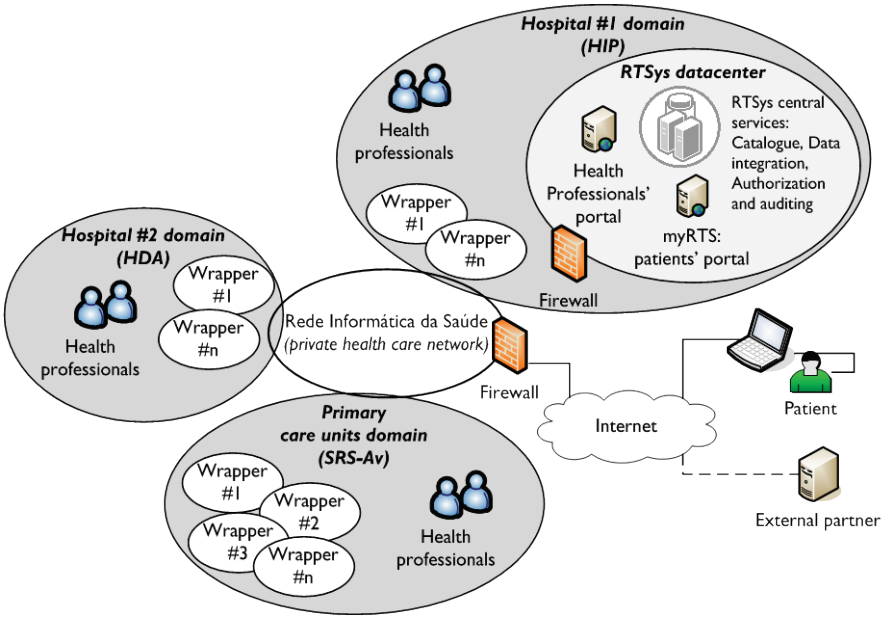


Figure 5.1: Deployment view of RTSys in the region of Aveiro, supporting collaborative workflows between two hospitals and six primary care centres.

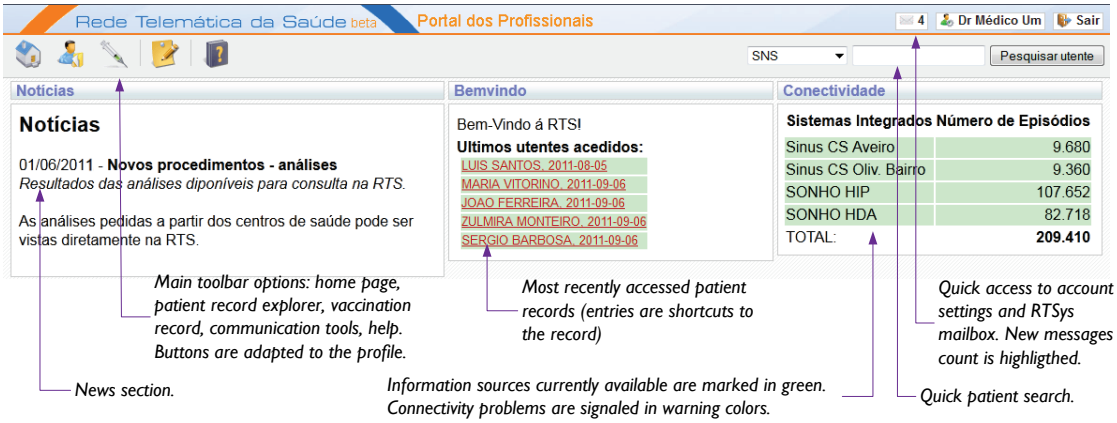


Figure 5.2: Health Professionals Portal: welcome page.

RTSys methods, and allows the health professionals at the partner institutions to navigate in the R-EHR. The eight institutions involved share the connectivity provided by RIS but, for the daily business, rely on independent health information systems and procedures. The initial deployment of RTS was performed according to the timeframe of the project, starting to pilot in January 2007. After that, several refinements and corrections were applied, as usual for software going live, originating a new version of the system.

The RTSys core services are deployed in the backend servers at the HIP data centre, containing the HIETA integration engine and the portals for end-user access. The wrappers (the data gateways) are installed in HIP and HDA, to access each hospital data sources, and at the SRS-Av data centre to access the primary care data. There is a wrapper for each source at each organization domain.

The health professionals access the RTSys functions using a friendly web portal, the Healthcare Professionals Portal, in which they can retrieve summaries for the patient encounters and interact with other professionals. The patient access to the myRTS portal from the Internet is possible since public connections are filtered through a firewall (at HIP). Although possible by design, external partners (*e.g.* private service providers) are not presently participating in the RTS.

5.1.1 PORTAL ENVIRONMENT FOR SEAMLESS ACCESS AT THE POINT OF CARE

The Health Professionals Portal (HP Portal) presents a friendly environment to navigate in the patient R-EHR at the point of care. The welcome page (after a successful login) provides access to announcements disseminated in the platform and shows the current availability of the distributed sources (Figure 5.2) ([5]). The user would start by looking up for a patient; s/he may wish to recall a recent case (and pick one of the recently used records in the list) or search for a given patient using the SNS number, for example.

[5] All patient data used in illustrations is fictional, including identifiers, such as the SNS number. The same holds for the information regarding professionals and sessions. The screenshots are provided in Portuguese, following the existing implementation.

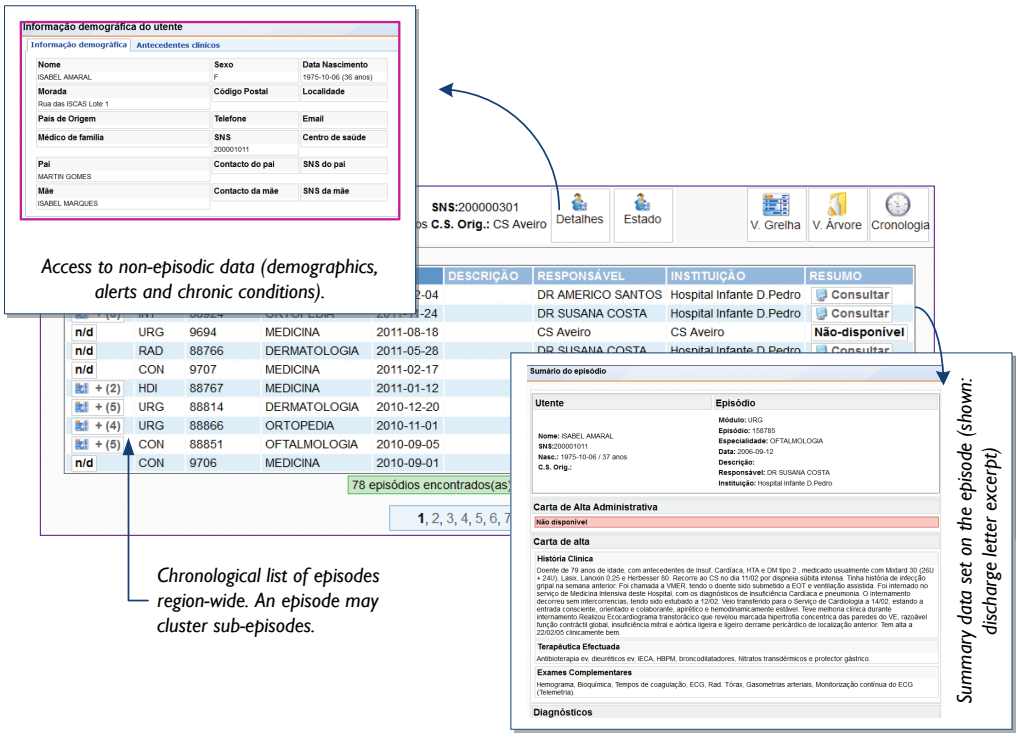


Figure 5.3: Health Professionals Portal: the R-EHR fed with encounters from several institutions.

Once a patient has been selected, his record, as indexed in the RTSys Catalogue, is promptly presented, structured as a list of encounters through time (Figure 5.3). From this table view, the health professional may choose to open the details of a particular episode, causing the retrieval of its summary data set from the remote source. The clinical alerts, detailed demographics and maybe other non episode-oriented data are available from the 'patient details' option. Episodes are classified after the familiar classes present in the main information systems in the region (SINUS and SONHO). Each class of episodes may originate a different summary but, for the same type, summaries are consistent, meaning that an in-stay discharge report produced in an Hospital, for example, would look the same as one produced in a different organization, providing a seamless access to remote fragments.

The public web site of the RTS project (http://www.rtsaude.pt/) makes available illustrative videos documenting the use of the HP Portal; the reader may refer to the multimedia section of the site for additional details on the supported interactions.

The Regional EHR visualization

The RTSys R-EHR includes encounters distributed in time and space, dimensions that should be clear to the end-user. The visualization of the EHR should adapt to the cognitive process of the health professionals as much as possible (Nygren *et al.*, 1992) and the organization of information is important for readability when accessing data from multiple sources (Wang *et al.*, 2010). The use of chronological views have been proposed in the literature to access the EHR (Plaisant *et al.*, 1998), sometimes complemented with knowledge base tools to provide problem-oriented visualizations (Bui *et al.*, 2007).

In RTSys, the default view to navigate in the encounters list is a grid view, sorted in reverse chronological order (Figure 5.4). The grid enables to condense several elements in a compact view and the health professional can quickly grasp the episodes over the time, their originating institutions and medical specialties involved. The grid layout can be interactively sorted (by clicking a column heading) to facilitate different ordering of the episodes (*e.g.* grouping by type, institution, etc).

An alternative presentation is the Tree view (Figure 5.4), a hierarchical organization of episodes according to their type. This facilitates, for example, the access to all episodes of a given type, for example, lab results, sorted chronologically.

A third layout available is the Timeline view, a sequence of events providing visual evidence of the episodes distributed in time. The time scale can be zoomed (condensing or expanding the period of time visualized) and events can be directly selected from the timeline to access details. It share commonalities with principles already suggested in (Plaisant *et al.*, 1998) towards a life line view of the health events.

This 'multidimensional cube' approach for the information layout allows the user to select

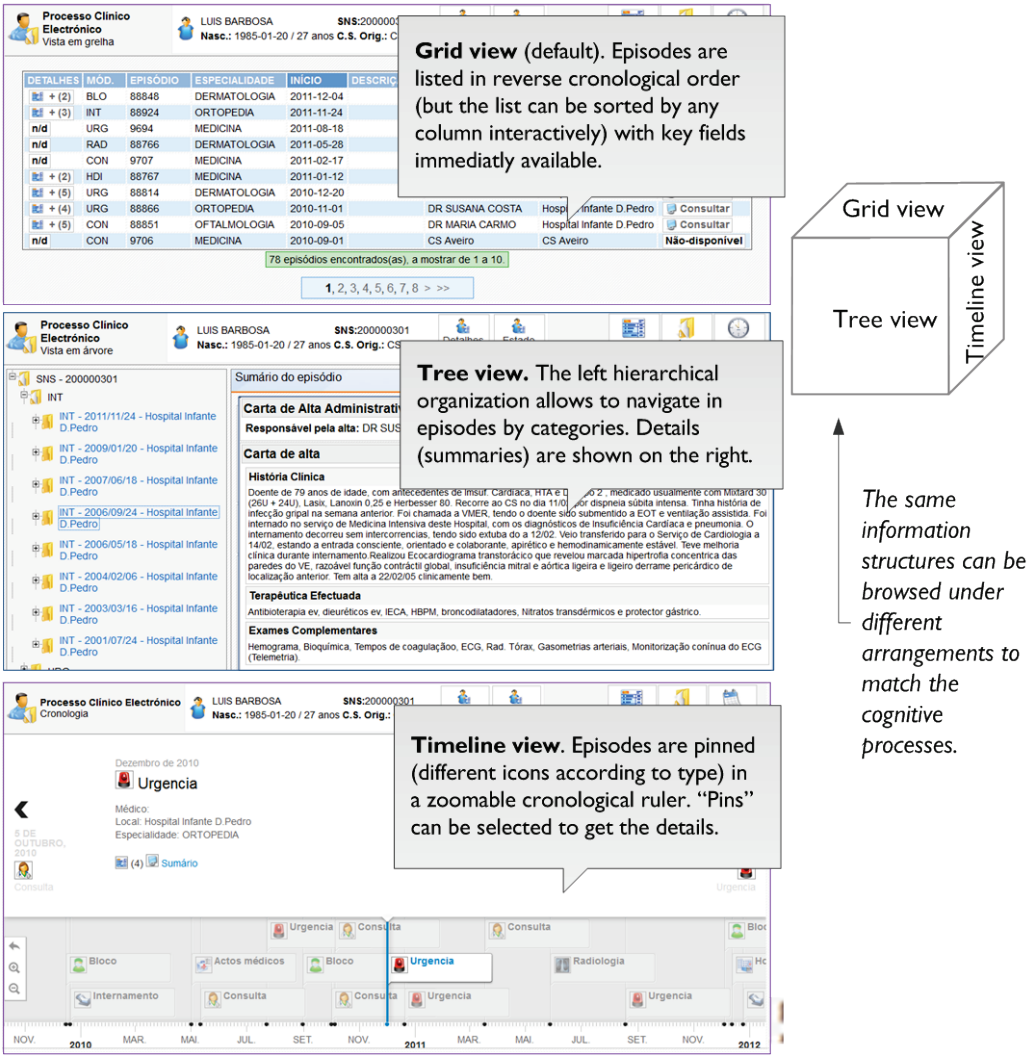


Figure 5.4: Alternative views to the Regional EHR.

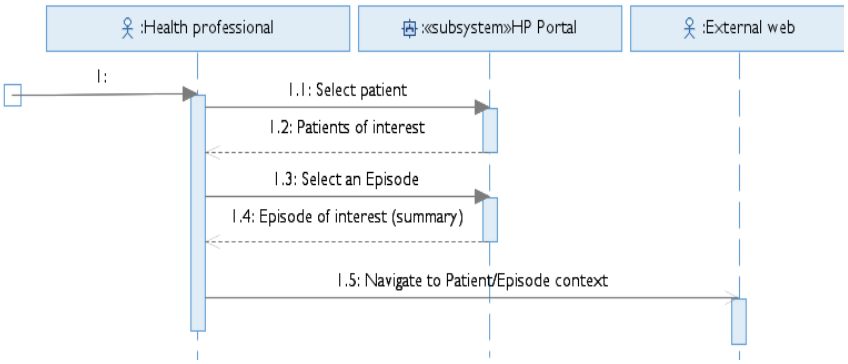


Figure 5.5: Control flow for web integration with external applications.

the perspective that best fits his cognitive processes. The views rely on the fact that the middleware is able to unify distributed data and present it as treatable entities through programming interfaces, facilitating different presentation strategies (unlike 'closed' binary blobs, for example).

Presentation layer integration

The presentation layer in RTSys is developed with Web technology and the information feeding the user experience comes from the HIETA middleware. The Web layer, by design, does not access directly the distributed, remote sources. There are cases in which it is not possible to access the external sources information structures programmatically. In these cases, it is not feasible to include the corresponding data in the HIETA discovery and integration processes. Some of these systems keep valuable details to document episodes available in the RTSys and already provide a dedicate web interface that can be used for presentation layer integration.

This approach has been deployed to integrate the user navigation in the HP Portal with remote web systems access. The redirection is initiated in the context of an on-going authorized session (in the HP Portal), in which the user has already selected a patient and episode of interest. The redirection URL used to invoke the target system encodes parameters required to allow that system to automatically recognize the RTSys user and set the usage context to match the episode or patient being accesses in the HP Portal. For this purpose, the codes for the Episode + Module (type of episode), or Sequential Patient Number are sent (these concepts come from the structure of the SAM-SONHO). Once redirected, the user has access to a different interfacing environment, in which RTSys has no control or knowledge of the actions performed (Figure 5.5).

In the Aveiro's RTS deployment, there are four external systems 'integrated' with the HP Portal using this approach. The linking into external applications is configured as much as possible to provide a seamless transition; an external application appears as a popup window of the RTSys (Figure 5.6). This redirection of the interface does not require the user to perform a new login at the external system. In fact, a basic single sign-on strategy is in place. This is a partial solution to work with the existing systems and not comparable to the functionality of a full-featured federated identity product, such as provided by the Shibboleth project (http://shibboleth.net), for example. In the RTS deployment, the external systems accept enough information in the URL to establish who the user is (a health professional), which patient or

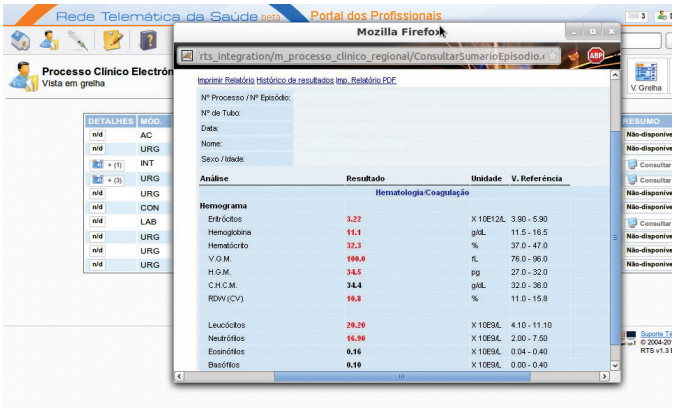


Figure 5.6: Example of contents from an external application displayed in the HP Portal session.

episode should be presented, and a security hash. In the case of the Appolo remote application (lab results), for example, the URL invocation is coded as:

```
http://<server>/appolo/wwwhidp/webappolo.exe/rts?chave=@appolokey&episodio=@episode&
modulo=@mod
```

The parameters (underlined) include a key, which is a hash using specific episode information (thus variable), agreed by RTS developers and the external system implementers.

The HP Portal has the ability to link to external web applications to facilitate, as far as possible, a seamless transition (preserving the patient context). The other way around is also relevant, *i.e.* the remote systems can also link to the RTSys portal. This is particularly relevant for the users of the SAM application, which is the solution that physicians working in public health care system mostly use (there are two editions, one for GPs and another for the physicians working in Hospitals). When the user is browsing a clinical record in SAM (the user has been authenticated and a patient is selected), s/he can invoke the option to open the HP Portal; the user intention would be, in fact, 'open **this** patient record in the HP Portal'. Passing the patient context is achieved by providing addition parameters in the redirection URL. In this case, the SAM request would present the following structure (parameters are underlined):

```
https://<HPPortal>/rts/m_processo_clinico_regional/ProcessaPesquisarUtente.do?scope=
&utenteIdType=2&utenteSns=snsNumber&cod_med=professionalId&is=organizationId
```

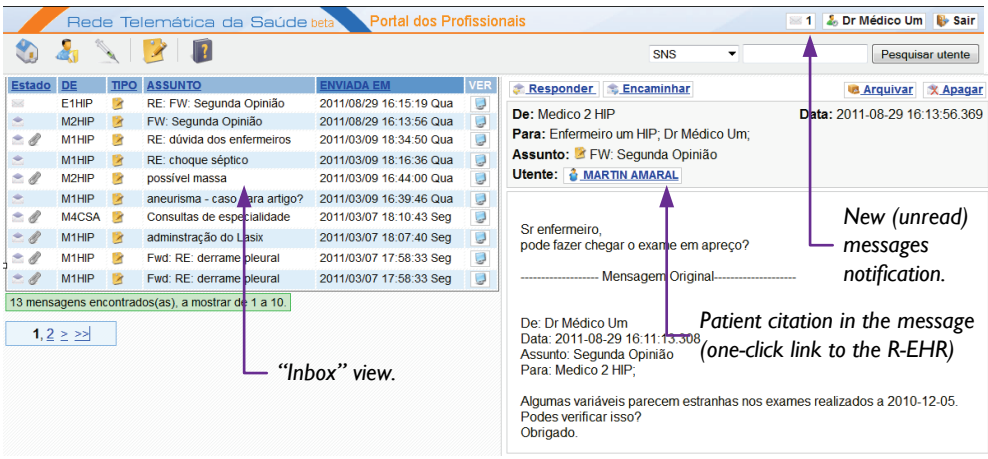


Figure 5.7: Secure messaging between RTSys users.

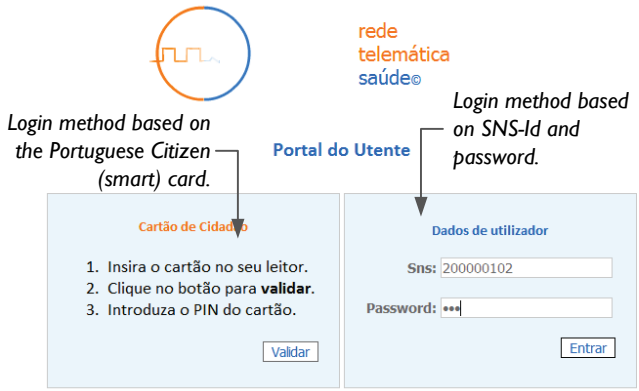


Figure 5.8: Alternative methods to login into myRTS.

For this specific integration scenario, no security token is exchanged and the user will have to do an explicit login in the HP Portal; the patient record, however, will be automatically selected in the Grid view perspective.

Communication and collaborations tools

The HP Portal supports the dissemination of profiled announcements (different news for different user roles) as a general communication and information tool. More interesting is the ability that users have to communicate with each other inside the portal. They can do this using a mailbox metaphor, similar to the user interactions familiar in e-mail clients (Figure 5.7). Messages lifecycle is managed internally by RTSys and they are not relayed across networks, as in the regular e-mail, providing a more secure communication facility between RTSys users. Besides the normal functions expected in a mailbox (read messages, compose, search recipients, etc.), the HP Portal supports the selection of patients to be linked in the message. This means that the message is about a given patient. When reading the message, the target health professional can open the R-EHR of that patient with a single click, promptly browsing the clinical context. This use case is known in RTSys as the 'patient discussion' (named after the discussions common in clinical practice without IT support). Another particular function to this messaging system is that it allows selecting target recipients from the health professionals known to RTSys, *i.e.* from the professionals' directory of the partner institutions (there is no need for email addresses).

5.1.2 PORTAL ENVIRONMENT FOR PATIENT PARTICIPATION

The patient-oriented use cases are supported in the myRTS portal environment. The user is any citizen in the region, with information in the underlying systems (the user must be registered in one of the primary care units in the consortium). The preferred authentication method is the use of the citizen card, a universal smartcard issued for Portuguese citizens (Figure 5.8). A username and password scheme is also available, as an alternative method.

The myRTS enables the citizen to browse a 'health agenda', *i.e.* the calendar of scheduled events in the partner institutions (Figure 5.9). A timeline of previous events is also accessible, in a calendar view. These events (past or future) are extracted from the production systems at the

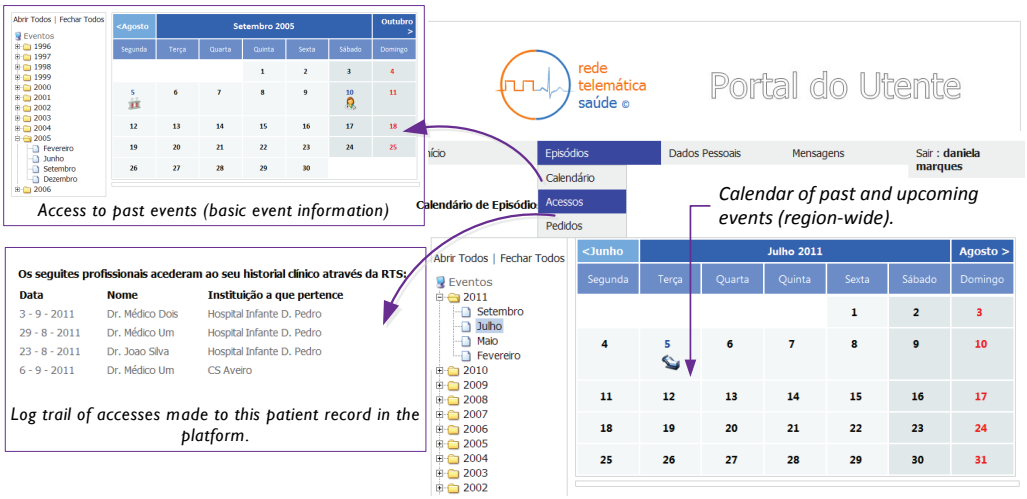


Figure 5.9: Overview of myRTS user interface.

partner institutions, especially from SAM-SINUS (primary care) and SAM-SONHO (Hospitals). Different icons signal different type of events, according to the medical specialities involved.

The patient can inspect the log of accesses made to his record in the platform (Figure 5.9). This list is extracted from the RTSys auditing database and presents the date and professionals that needed to access the record.

Basic messaging is supported, enabling the patient to address messages and simple requests to his GP. Both will be available in the physician mailbox, accessible through the HP Portal. This information, once submitted, is securely transported inside the RTSys platform. The GP may reply using the counterpart functions in the HP Portal.

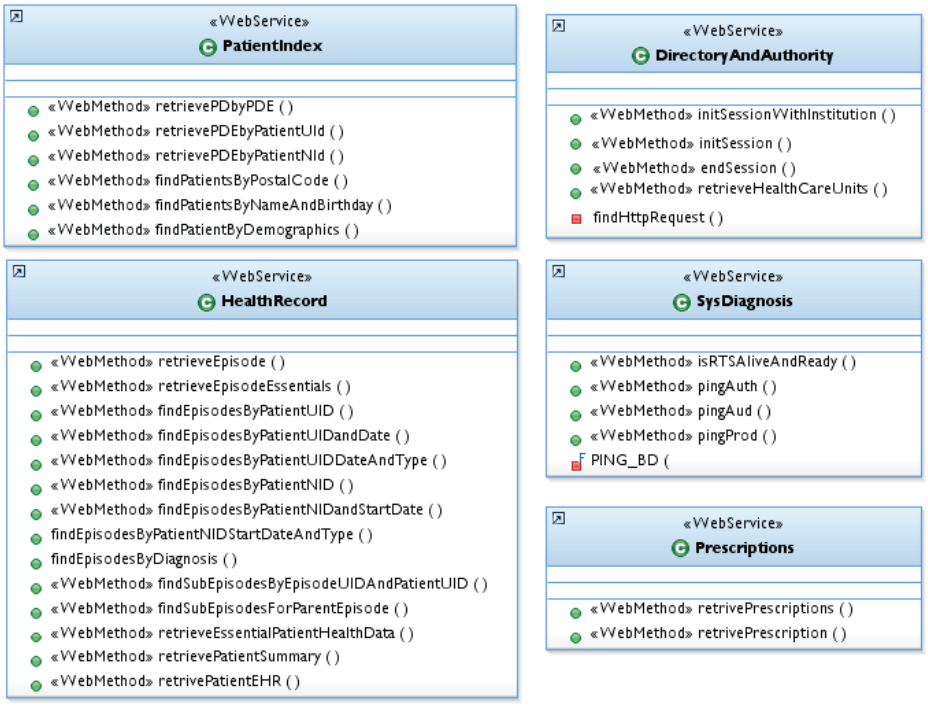


Figure 5.10: Main web methods implemented in the RTSys services.

5.1.3 SERVICES FOR APLICATIONS INTERFACING

The RTSys offers a set of services that allow third-party applications to implement new scenarios, exploring the regional scope (Figure 5.10). The services are provided as secure Web Services, implementing the WS-Security standard (WHO, 2008). The client applications need to supply a valid digital certificate to initiate the services invocation (which must be requested to the Consortium). This provides a sound control over the applications that can connect. A second level of authorization occurs to identify the user and associated profiles. The application must supply user credentials to the *initSession* method which verifies the access control lists and, if the user is recognized, returns a session token that must be supplied in subsequent service calls. The actions of client applications must have an associated user.

Data structures returned by the services distinguish between basic and a complete versions of the main entity. For example, the basic demographic elements form the 'patient demographics essentials' (PDE) while the complete version is just 'patient demographics'. This allows having web methods that just return the 'essential' version of the entities, and other to get the full content, optimizing the network interactions.

As a proof of concept, we implemented a demonstration client that gets the information available thought the RTSys and produces a patient summary compatible with the epSOS project specifications.

5.2 System interactions for data discovery and integration

The source systems in RTSys are completely independent from the middleware operations and are not required to be adapted in any way to participate in the integration platform. They should, however, provide some mechanism for technical interfacing with external applications, required by the RTSys wrapper module. The middleware will visit the registered systems, as configured in the deployment configuration, to discover patients and new care episodes. The collection of source systems can vary dynamically, since the RTSys methods allow for changes on the number and availability of the source systems.

The key processes involved in the dynamic data discovery and access (presented in the following sections) are (1) the acquisition of new Patient entries, (2) discovery of distributed episode data and updates, and (3) the on-demand distributed data access.

5.2.1 SYSTEM INTERACTIONS TO BUILD THE MASTER PATIENT INDEX

The Master Patient Index (MPIx) or Patient Registry is a well known component in distributed health care information systems in which it is necessary some source of truth regarding the patient identification. This is the function of MPIx in HIETA, the component responsible for keeping patient demographics, answering simple patient lookup queries and applying data consistency checks for patient identity management.

At the time when the development of HIETA started, there was no single patient registry solution for the National health system. That has been changed with the introduction of the *Registo Nacional do Utente* (RNU). Being the official source of patient identification, RNU provides a reliable source of patient identification. The problem with RNU is that it is not always reachable by the RTSys processes, due to several network service-level fluctuations.

All partner organizations using RTSys already have information systems that identify patients. This means that patient demographics are available from multiple sources, but not necessarily consistent across them. The MPIx addresses this problem ranking the information sources, favouring the demographics from the source with the higher rank (among those in which the intended patient is available). The ranking is based on simple heuristic: in first place, the national reference patient registry system is accessed (RNU), which is known to apply data curing procedures; if not available, then the source information system representing the primary care unit in charge of the patient is used; if this wrapper is not available, then any source can contribute with the patient demographics.

The MPIx provides another important function which is 'identity brokering' capabilities, establishing the correspondence between the identity numbers of the patients at the different data sources. Tracking identities at different sources is facilitated by the use of a shared key already disseminated in the target deployment, the national health system number (SNS number).

Finding patients

To find a patient, a user can search using the following parameters: the national health system identity number (SNS-Id); a combination of name and gender; a combination of name and birth date; or the RTSys internal identifier. (The latter option is occasional used, since the users are not exposed to the internal RTSys identifier nor are they expected to learn it.) The most popular option is searching by SNS-Id, a practice that is common in other information systems existing in the partner institutions.

When the patient is searched in a client application, *e.g.* a portal, his record may already exist in the MPIx and HIETA will return the corresponding demographics. The user may them continue the data retrieval use case, probably asking for details on available encounters. The acquisition of new patient demographics occurs when the patient data is effectively used, *i.e.* when the user lookups up for that patient (Figure 5.11). If the patient has not been accessed before, HIETA detects a 'patient miss' (somewhat similar to the page-fault event, in the computer architecture domain (Watch, 2005)) and triggers the distributed lookup, forwarding the query to the (wrappers at the remote) information sources, following the ranking strategy previously presented. Once an entry exists in the RTSys Catalogue, the MPIx resolves patient lookups without pooling the remote information sources (Figure 5.12).

MPIx entries lifecycle

In a new deployment of RTSys, the Catalogue (which includes the MPIx) would start empty. New patients are learned by the RTSys index as they are effectively used. This strategy can be

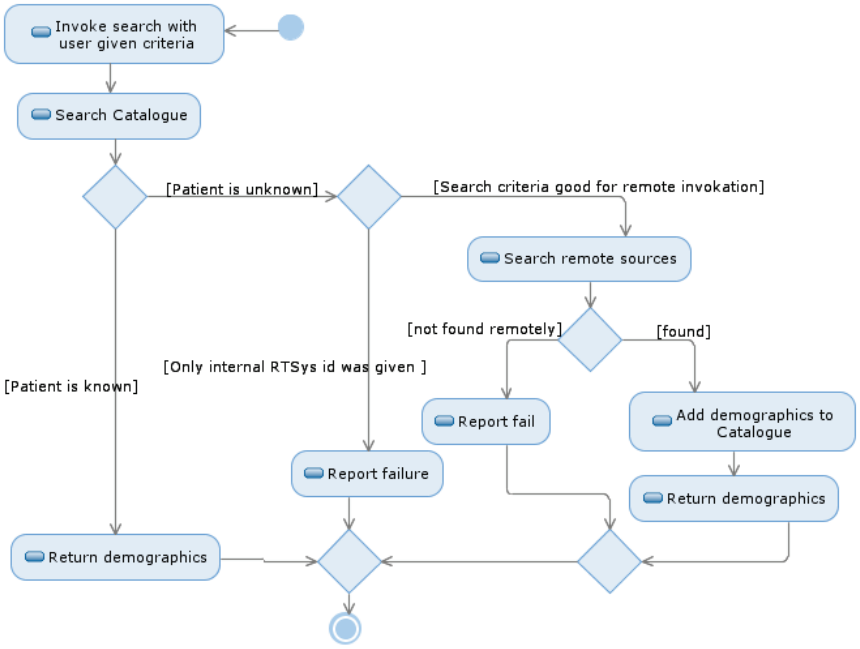


Figure 5.11: Patient lookup flow.

SNS	NOME	GÉNERO	DATA NASC.	CONSULTAR PROCESSO CLÍNICO			Boletins
200000478	LUIS TAVARES	m	1974-03-19 (38 anos)	V. Grelha	V. Árvore	Cronologia	Boletim de vacinas
200000308	LUIS VITORINO	m	2003-05-04 (9 anos)	V. Grelha	V. Árvore	Cronologia	Boletim de vacinas
200000105	LUIS SANTOS	m	1975-04-12 (37 anos)	V. Grelha	V. Árvore	Cronologia	Boletim de vacinas
200000008	LUIS MONTEIRO	m	1982-02-13 (30 anos)	V. Grelha	V. Árvore	Cronologia	Boletim de vacinas

4 utentes encontrados(as), a mostrar todos(as) os(as) utentes.

Rows tagged with a yellow bullet correspond to patients (available in remote systems) not yet indexed in RTSys.

Figure 5.12: Patient lookup results (basic identification attributes).

complemented with other proactive behaviours, for example, using the schedule appointments (*e.g.* for the next day) to autonomously start patient integration cycles.

A supporting assumption is that patient demographics are reasonably stable and the data in the RTSys Catalogue recently obtained is good for subsequent 'hits'. There are, however, two ways to invalidate this data, originating the re-collection of an up to date version from source systems. The first mechanism is a user driven refresh option, by which the user can force a full update of the patient data. This option is available to compensate for eventual problems that may have occurred when the patient was previously integrated (*e.g.* the RNU was not available at the time), but is unlikely to be used often. The normal consistency verification method is a simple time to live (TTL) tagging mechanism. This sets a period of time in which the MPIx entry is considered good, this is, that the corresponding data is assumed to be up to date. After that time, the entry is invalidated in the MPIx (tagged as 'dirty', but not deleted). If the information on that Patient is then requested, it forces a full patient integration, recollecting the data from the source systems. The TTL parameter is set by configuration and can be used to control the updating of the RTSys Catalogue.

5.2.2 SYSTEM INTERACTIONS TO BUILD THE REGIONAL PATIENT RECORD

The data federation implemented in HIETA is based on two main components: (1) distributed wrappers abstract each information source and expose a common services interface and information model; (2) the central planner invokes the services on the wrappers to fulfil periodic or demand driven integration plans. In this sense, RTSys implements a pull model for information extraction, transformation and loading, since it is up to the central components to ask providers for data (and not the wrappers publishing changes, another common architectural option).

HIETA indexes the existing known care episodes for a patient in the Catalogue, along with the source location of each data fragment (where to get the details). The use of the index enhances the overall performance since it allows for fast lookups in the central index and, when required, on-demand access to the source systems to obtain the episode details. Simple queries, such as searching for a patient and retrieving a list of encounters, are promptly satisfied by the information in the Catalogue.

The processes described for the acquisition of new patient demographics are, in fact, embedded in a larger process, which is the gatherer of the minimal information for the RTSys index entry (including demographics, a list of known care episodes and basic clinical alerts, if available). Data discovery will ensure that the index is able to (1) acquire new patients and entries to the regional record, and later (2) update the information for existing patients. These two processes imply different object collaborations, as described below.

Integration of a new patient record

A 'full patient integration' is the process of collecting all information on a Patient, by visiting all the information sources. This process will take place in two different cases: when the patient is not yet indexed in the Catalogue, which means that it has not been used before (the information on the patient has never been requested); or when it is ordered on-demand. Starting with the later, the full patient integration (FPI) can be asked in the Professionals' Portal if the user, for some reason, wants to force the reintegration of a patient case (as introduced previously for the MPIx). Note that the FPI will also imply that the previous information in the Catalogue is deleted and, in practice, this acts as a user-driven full refresh mechanism of the record (the user, however, is not expected to do this on a regular basis, as an automatically update mechanism exists). The typical activation of a FPI takes place when a 'cache miss' occurs, *i.e.* an attempt was made to access the patient information in the Catalogue but there is not yet information about that patient (or the record exists but is tagged as 'dirty').

Incremental patient data updates

HIETA tracks updates to the patient data by pooling the sources for updates, according to a configurable frequency. This process is typically deferred to idle times, during the night, according to the deployment configuration files. The periodic patient integration (PPI) comprises three main phases: verification of the available sources, request for episodes from each source, and integration in the Catalogue. Updates detection is supported by using a provenance trail to

Listing 1: Overall algorithm for periodic episodes updating. Details were omitted for the sake of readability (e.g. exceptions handling, logging, etc.).

```
    // which sources are ready?
    List readySourcesList = MonitoringToolkit.getReadySources();

    // create thread resources for parallel data retrieval
    for (Source source : readySourcesList) {
        WrapperClient client = SourceWrappersFactory.proxyFor(source);
        threadHashmap.put(source, new ThreadedWrapperClient(client));
    }

    // update the information for the patients in the Catalog
    List<Patient> patientsList = Catalog.getListOfActivePatients();
    for (Patient patient : patientsList) {
        for (Source source : readySourcesList) {
                // get the id of this patient in the remote system
                String remoteId = MasterPatientIndex.resolveLocalIdentity(
                                patient, source);
                // get the last good integration for this source and this
                // patient
                Date lastGoodintegration = Catalog.getLastGoodintegration(
                                patient, source);

                // set the query parameters
                WrapperClient client = threadHashmap.get(source).getWorker();
                client.setParameters(patient, remoteId, lastGoodintegration);
                // starts the remote invocation
                threadHashmap.get(source).run();
        }

        // harvest the results per source
        for (Source source : readySourcesList) {
                // wait for the thread to complete
                threadHashmap.get(source).join(MAX_WAITING_THRESHOLD);
                // the client object has the data payload
                WrapperClient client = threadHashmap.get(source).getWorker();

                Catalog.update(patient, source, client.getResults(),
                                currentTime());
        }
}
```

tag episodes with basic metadata, including the last known good integration timestamp. Only the episodes occurring after that time are required to be fetched.

The overall algorithm is displayed in Listing 1. We chose to adopt the Java syntax, for clear semantics and familiar abstractions; in practice, additional code is required, such as objects declaration, initialization and exception handling, which we have suppressed for legibility sake.

5.2.3 ON-DEMAND DATA ACCESS

The previous processes ensure that new patient records are assembled and kept updated in the platform. To perform their functions, they rely upon (and update) the RTSys Catalogue. Later on, the user requests in the Portal originate the following sequence of operations (Figure 5.13):

1) The regional record for a given patient is requested in the portal. The request is routed to the Integration Engine internal module.

2) The Catalogue is inspected for the list of episodes on that patient. If no patient is found at all, then a data miss exception is generated; the user will be asked if s/he wishes to order a live (on-demand) patient search in the distributed sources. If, instead, the record

is found, the patient demographics, along with the list of episodes are retrieved from the RTSys Catalogue.

3) The user explores the list of episodes (*e.g.* sorting, filter by type, etc.) and may identify an encounter of interest for which s/he may need additional information, selecting the corresponding option in the portal.

4) The request for the episode details is routed to the integration engine. The episode data comes from a specific information source, abstracted by an RTSys Wrapper. The integration engine prepares the query for that Wrapper, expressed in Object Query Language . The remote Wrapper is invoked using a Web Services over SOAP.

5) The Wrapper maps the request, expressed in the global information model, in a specific query for that source. The use of libraries for object-relational mapping facilitates the query transformation and, in RTSys, that function is fulfilled by the Hibernate framework[6].

6) The results from the source are mapped in the global model by the Wrapper, marshaled in XML and sent back to the Integration Engine module.

7) The Integration engine releases the data to the presentation layer, in this case, the portal environment, and the health professionals get the remote information formatted consistently.

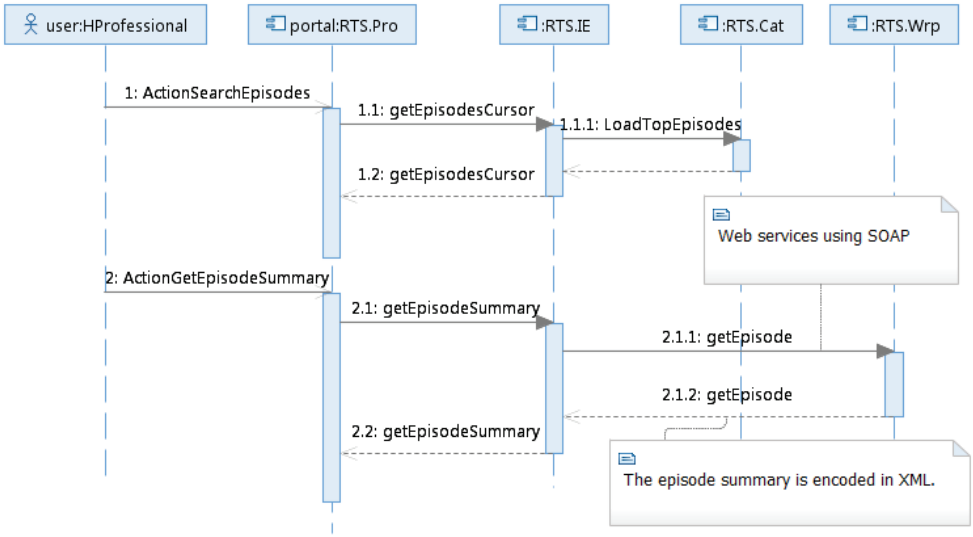


Figure 5.13: Episodes metadata is kept in the catalogue; the remote wrappers retrieve the actual summary.

5.3 System monitoring and auditing

The RTSys deployment in Aveiro uses the existing networking infrastructure for the public health organizations, which supports a great variety of systems and applications, over different

[6] Hibernate – JBoss Community. Available from: http://www.hibernate.org/

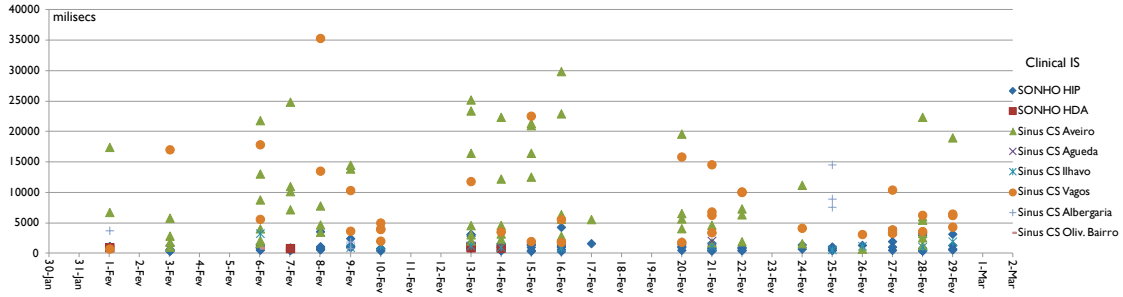


Figure 5.14: Performance for the several sources in RTS deployment (excerpt for Fev-2012).

bandwidth links. The network availability can affect the service level of RTSys in multiple ways and motivated us to develop specific monitoring tools.

Besides the normal IT infrastructure observation, using the Nagios[7] system to produce reports and alarms, a specific monitoring strategy was developed in RTSys to verify the performance of critical services, namely the distributed integration process. This monitoring toolkit tracks the evolution of the following variables: local information sources data retrieval time, data transfer times, single patient data integration time, and overall Catalogue update process. The monitoring toolkit was presented in a related publication (Santos *et al.*, 2009).

The monitoring tools allow for the visual inspection of bottlenecks. Figure 5.14 plots, for each remote clinical information system, the time required to integrate a patient (*i.e.* get the patient data and all episodes from that source). The different plots for the same source in the same day correspond to distinct patients. The concentration of plots is under the five seconds threshold (the integration of a patient record at that source took less than that), but outliners are frequent and correlate with primary care units servers (remember that the central services of RTS are installed in the large hospital).

Auditing

The audit trail kept by RTSys records each action performed by authorized users in the platform. For the purpose of later inspection, certain actions, such as opening the patient record, are described with key parameters, allowing clarifying, for example, which specific patient was accessed. The audit trail is available in the supervisor dashboard but also shown, as an excerpt, in the health professional and patient portals (Figure 5.15). The supervisor can fully inspect the audit trail, filtering for a specific user and period; technical details are included (*e.g.* the requesting IP address) as well as the parameters used in each request. The health professional will have a small window listing the most recent accesses s/he made; besides a convenient way to re-open a patient case, it also assures the user that logs are collected. For the citizen, the excerpt includes a list of accesses concerning his own record, clarifying the time and health professional that opened the record (in the RTSys platform).

[7] Nagios IT Infrastructure Monitoring. http://www.nagios.org/

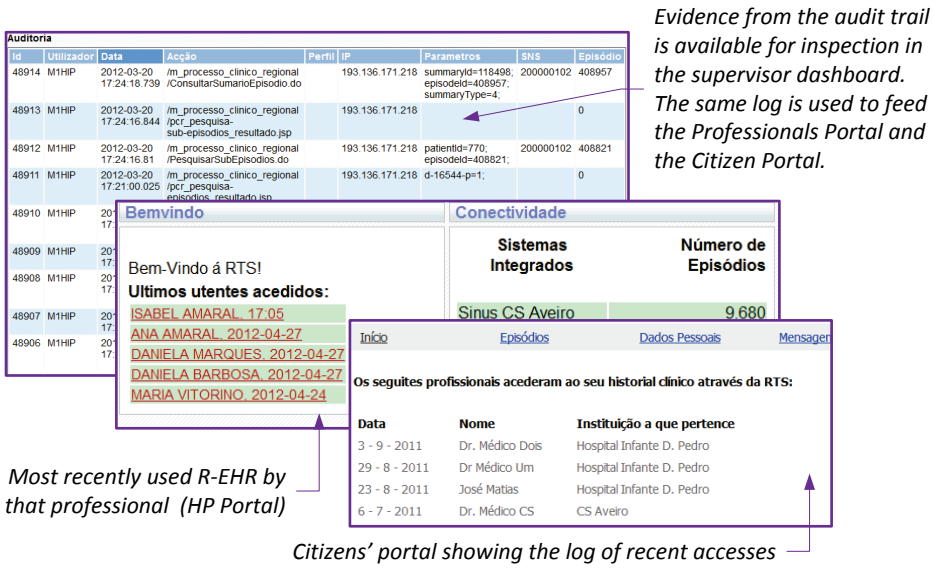


Figure 5.15: Excerpt of the auditing trail shared by multiple actors.

5.4 Software stack and development practices

For the developer, the RTSys solution is multi-module Maven[8] project, written in Java, according to the J2EE paradigm (Alur *et al.*, 2003). The domain logic is deployed as components that run at web containers (the Apache Tomcat[9]). The development practices take advantage of the Maven *plug-in* structure to orchestrate continuous integration to provide automated quality assurance checks.

5.4.1 SOFTWARE PROJECT MODULES

The internal architecture of the RTSys software comprises several modules with complementary functional scopes (Figure 5.16). The concept of a module can have different interpretations but, in this case, we keep the semantics of the Project Object Model (POM) used by the Maven build-automation tool. A module is a software component that can be individually packaged and managed. The Maven tool facilitates the declaration of dependencies between modules, either internally developed or obtained from a structured world-wide library or modules, automating several aspects of the build process.

The purpose of each module is summarized in Table 5.1. The access layer includes the portals (RTS.PRO, RTS.UTE) and the services interface (RTS.WS); they all rely on the entry points provide by the integration engine (RTS.IE), which acts as the business logic coordinator. Another key module handles the directory of service points and actors, and the authentication and authorization functions (RTS.DA). Several components are available dealing with data semantics and persistence. The key metadata is kept at the RTSys catalogue (RTS.CAT), which also

[8] Apache Maven project. Available from: http://maven.apache.org/
[9] Apache Tomcat. Available from: http://tomcat.apache.org/

Table 5.1: Brief description of the RTSys internal components.

Component	Description
RTS.AUD	Auditing functions (store and access the audit trail).
RTS.CAT	RTSys Catalog management, including metadata queries and the persistence of the index.
RTS.DA	Directory and Authorization, responsible for keeping the professionals directory and access policies, and resolve authentication and authorization challenges.
RTS.DTS	Data Transformation Services, responsible for converting and verifying remote data.
RTS.IE	Integration Engine; provides the entry point to all data requests and initiates the required distributed interactions.
RTS.MPIx	Master Patient Index; manages the directory of patients known to the RTSys.
RTS.PRO	The Health Professionals portal (web application).
RTS.RIM	The reference object model, establishing the "vocabulary" of the RTSys.
RTS.UTE	The myRTS portal (web application). The acronym comes from the word *Utente*.
RTS.WRP	The wrappers; each specific wrapper is implemented by extending a basic RTSys wrapper.
RTS.WS	The Web-Service API for RTSys (over SOAP).

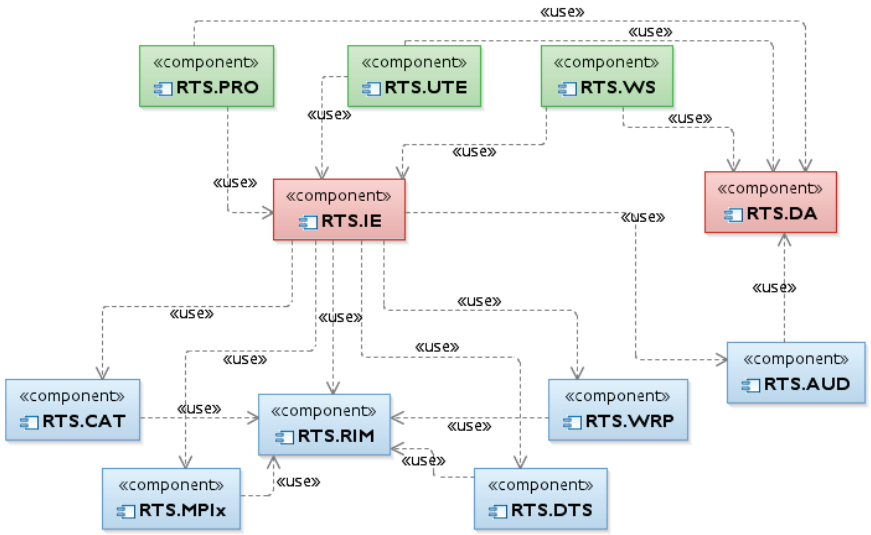


Figure 5.16: RTSys internal components and main dependencies.

implements the health record data consistency logic. The Catalogue is complemented with the reference patient registry (RTS.MPIx). Data transformation (in the central services) is handled by RTS.DTS, while the remote data adaptors belong to the Wrapper module (RTS.WRP). The common information objects are shared in the reference information model (RTS.RIM), a pivotal component.

5.4.2 SOFTWARE CONSTRUCTION ENVIRONMENT

Software quality assurance is a complex and multi-layer problem (Lewis *et al.*, 2008). Part of the risks can be mitigated by a sound construction process (McConnell, 2004). In the development

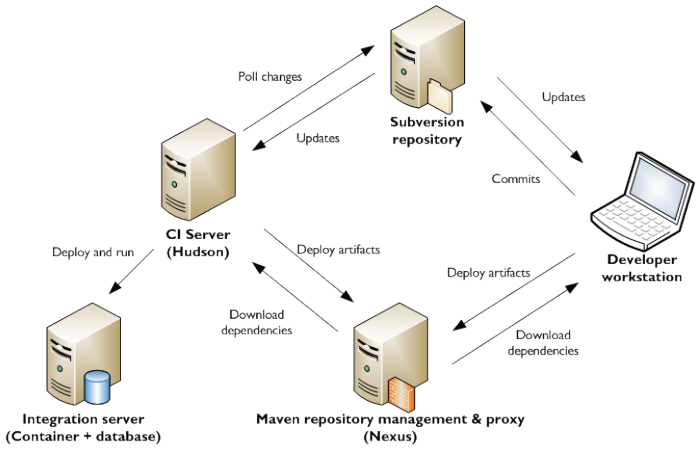


Figure 5.17: The continuous integration interactions.

of RTSys we adopted a set of practices to ensure that software defects would be detected early in the development process, by adopting an automated continuous integration approach (Paul, 2007).

In the RTSys development setup, the 'production line' includes several actors and interactions (Figure 5.17): the developers work at their workstations, implementing a piece of functionality previously assigned. Before adding new functions, the developer should get the most up to date version of the code base from the team Version Control Systems (VCS), in this case, the Subversion[10] server. When a small-grained, but relevant piece of functionality is completed, the developer ensures that the solution has no syntax errors, tests it (white-box unit tests) and commits the changes to the VCS. The continuous integration (CI) server, supported in Hudson[11], will detect new changes in the VCS and starts a code test cycle. These tests may include module integration verification, requiring the CI server to do a staging deployment of the solution in a specific integration environment to test that new modules (database, applications in enterprise containers, etc.) work together. The Maven build tool enables the introduction of advanced features in the build process, such as the use of a local repository of binaries, from which dependencies can be downloaded and new artifacts can be made available. The repository is available in the local network and managed by the Nexus[12] solution which also supports acting as a proxy to global repositories available elsewhere in the internet.

The construction line of RTSys enables a productive environment, with early detection of defects (as long as the developers commit often). The central piece in this process is the automated work from the CI server (Figure 5.18):

1) The CI server periodically pools the VCS for changes. If new code is found, the CI updates its local version of the project to test it.

2) The Maven project configuration file (Project Object Model - POM) is read and all the dependencies between modules and from external libraries are identified. The

[10] Apache Subversion. Available from: http://subversion.apache.org/
[11] Hudson Continuous Integration. Available from: http://www.hudson-ci.org/
[12] Sonatype Nexus. Available from: http://www.sonatype.org/nexus

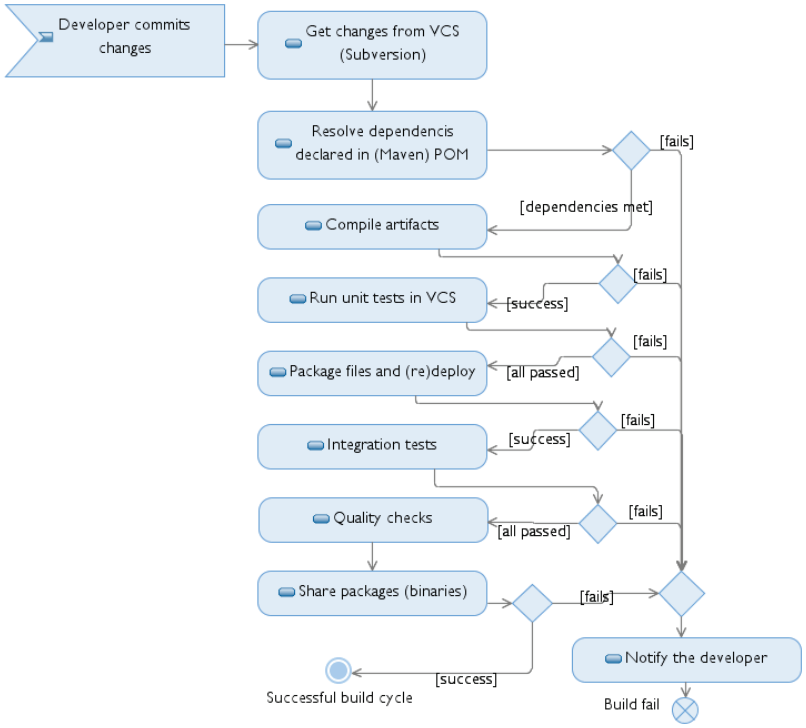


Figure 5.18: The integration cycle executed by the continuous integration server.

dependencies must be downloaded and here the use of a local repository management (Nexus) is a key element to speed up binaries transfer.

3) With the code and all required dependencies locally available, the CI server compiles the source code and related resources. This is the updated, all-team contributed version of the project.

4) The CI looks for unit tests available in the source code. The tests are implemented using the JUnit[13] library and automatically recognized by the Maven tool. If found, all of them are executed (or until the first one breaks).

5) The next step is the packaging of the compiled modules, resources and configuration files into purpose specific assets. The deployment for the Apache Tomcat container is a web archive format, for example.

6) Again, the CI server will ensure the redeploy of artifacts in the application servers (of the staging environment), with the Cargo[14] Maven plug-in. The setup may require that the databases are migrated to present schemas consistent with the updated development release. For this purpose, the Liquibase[15] Maven plug-in is used.

[13] JUnit.org. Available from: http://www.junit.org/
[14] Cargo. Available from: http://cargo.codehaus.org/
[15] Liquibase. Available from: http://www.liquibase.org/manual/maven

7) The integration tests are executed in the staging environment. These tests include programmed actions to verify multi-class logic, as the services supported by RTSys web services (using soapUI[16] tool).

8) When the automated unit and integration tests are completed, the CI will start a static analysis of the code quality, looking for error patterns. This is ensured by quality plug-ins loaded into the CI server, *e.g.* FindBugs[17].

9) If the tests are successful, the new assets (build results) are uploaded into the shared Maven repository and available to the team.

The sequential test cycle at the CI is mapped in the Maven build lifecycle, which is indeed the workhorse behind this automation. The test cycle breaks as soon as a problem is found; integration tests, for example, will not take place if the unit tests do not pass. In case of a failure, the responsible developer is promptly notified with the feedback from the build process. If all checks pass, then new binaries are uploaded to the shared repository, free of unit level errors, integration problems and problematic code patterns. The acceptance tests and system validation, however, are still a people factor. Other high level project management practices, such as the agile approach supported by Scrum, are not covered by this construction process explanation.

5.4.3 FREE SOFTWARE STACK

The implementation of the RTSys distributed system adopts the Java Enterprise reference model (J2EE) and associated standards, which are particularly suited for enterprise integration (Alur *et al.*, 2003). The J2EE architecture allows, in particular, standards based deployment of web and application servers, and object-to-relational data management. These advanced features provide a sound basis to deploy scalable solutions since it makes feasible the exchange of components in the architecture, for example, adding more application servers in the business logic tier, or changing the persistence provider to a higher performance library or database server.

The use of a reference deployment architecture such as J2EE, with several alternative implementations of its components (including public-domain resources), favours the use of free software. This is the case of the RTSys (Table 5.2): databases required for the central services, the Catalogue and Auditing, are supported in PostgreSQL[18]; persistence takes advantage of the Hibernate[19] library to supply the object-to-relational framework; the *servelet* container is the Apache Tomcat[20] and the counterpart Apache HTTP server[21] is used for the web server component and deployment of reverse proxy policies. The web layer is also supported on free

[16] SmartBear soapUI. Available from: http://www.soapui.org/
[17] FindBugs. Available from: http://code.google.com/p/findbugs
[18] PostgreSQL open-source database. Available from: http://www.postgresql.org/
[19] JBoss Hibernate project. Available from: http://www.hibernate.org/
[20] Apache Tomcat. Available from: http://tomcat.apache.org/
[21] Apache HTTP server (httpd). Available from: http://httpd.apache.org/

libraries, including the Model-View-Controller (Gamma *et al.*, 1995) framework from Apache Struts[22].

The present choice of technologies allows for a deployment free of licensing costs to the health care organizations which, given the financial constraints of health care systems, especially in the present economic context, seems to be a relevant argument. Besides being free, the major components in the software stack are also open source, which means they can be inspected by the developer community for eventual functional or security flaws.

Table 5.2: The RTSys software stack.

Level	**Technology and purpose**
Presentation layer	**JavaServer Pages (JSP)** – The technology used to serve dynamic pages. **JavaServer Pages Standard Tag Library (JSTL)** – Extension to the JSP that includes tags adding support for importing pages, creating logical conditions, creation of parameterized URLs, processing arrays, etc. **JavaScript** - Prototype-based scripting language used to improve the client interface. **Apache Struts** – An open-source web application framework that separates the application logic from the view. **Apache Tiles** - Templating framework used to simplify the development of the user interface. **DisplayTag** - The display tag library is an open source suite of custom tags and is used in the creation of dynamic tables all over the RTS. **Jenkov Prize Tags** - Collection of tag libraries for internet/intranet applications used to create the tree view of the episode list.
Application containers and server environment	**Apache HTTP Server** – The HTTP server used. **Apache Tomcat** – The container that contains the applications and web services. **Apache Axis** – The web service stack responsible for the services that access the external systems. **Apache log4j** – The logging system used by all the artefacts. **Metro** - The web service stack responsible for the RTS web service layer. **Quartz Scheduler** - Job scheduling service responsible for controlling the scheduling of updates and notification sending. **J2EE** – Java enterprise edition specifications and reference APIs.
Persistence layer	**Hibernate**. Object-to-relational mapping library. **JDBC**. Universal driver for relational data access in Java (combined with Hibernate). **PostgreSQL** –Open source object-relational database system used to store the RTS Catalogue and other databases.
Operating system	**Linux OS (CentOS)** – Servers are deployed over Linux (CentOS edition).

[22] Apache Struts. Available from: http://struts.apache.org/

6 Results and deployed prototypes

The RTSys telematic platform for regional connected care was deployed in the context of the RTS project, in the region of Aveiro. In the present chapter, we describe this specific deployment, the applications available to the end users and the pilot use of the system.

6.1 Results

RTSys was an answer of our research group to the opportunity raised by the *CidadesDigitais* programme to enhance the availability of telematic services for the society. Building on previous research in the group, and working with the domain experts that participated in the specification activities, we designed and implemented the RTSys, a new platform to enable care organizations to connect and share clinical data with partners, in a controlled and secure environment. To our best knowledge, RTSys was the first solution in Portugal, developed with the participation of the health professionals in the field, aiming at a generic federation system to build the virtual regional EHR across organizations.

The RTSys instance deployed in the region of Aveiro, within the RTS Project, is known as just RTS. This setup is registered with the Portuguese data protection agency and has the required clearance to operate.

6.1.1 THE RTS DEPLOYMENT

RTS connects the two major hospitals in the region of Aveiro and six Primary care centres since early 2007, covering about ~350,000 citizens. Each primary care centre refers patients to one of the Hospitals, and the Hospital in Águeda refers patients to the Hospital in Aveiro (recall Figure 1.4, p. 6). RTS integrates the core information systems in these institutions (keeping track of patient admissions and episodes), allowing to build a comprehensive view of the encounters that took place in the region. The clinical information on these patients forms a universe of ~19 million episodes registered along the eight institutions (Table 6.1).

All partner institutions belong to the public National Health Service, *Serviço Nacional de Saúde* (SNS), and are connected through a private data network, *Rede Informática da Saúde*, administered by the SNS. This infrastructure is exclusive for care providers, which facilitates the connectivity without additional costs and ensure an increased level of security. The computing resources supporting RTS are hosted inside this private network, at the larger Hospital (HIP)

Table 6.1: Data volume by partner institution (approximated values for Nov-2011).

Partner institution	Patients [a]	Episodes [b]
Hospital Infante D. Pedro	410.400	6.646.400
Hospital Distrital Águeda	125.900	1.557.300
PCU - Águeda	73.700	2.459.300
PCU - Albergaria	50.300	1.074.500
PCU - Aveiro	135.300	2.948.800
PCU - Ílhavo	83.300	1.860.800
PCU - Oliveira Do Bairro	43.500	1.360.700
PCU – Vagos	39.100	1.114.700

a) Patients registered in the Institution principal Admissions information system.

b) All episodes, without the RTSys clustering/hierarchy.

datacenter, benefiting from the existing technical facilities. The internet access to specific services, such as the myRTS portal, is filtered by the firewall at HIP. External partners (*e.g.* institutions supplying lab analyses) may also connect to participate in cooperative workflows by design, but not present in the current production environment.

Institutional contacts to expand RTS to other public care organizations were established. The inclusion of new partners and information systems is supported by the platform, but the required organizational actions did not take place and RTS keeps the initial consortium.

Each care organization contributes to the R-EHR with different information sources. The two hospitals use the same patient management solution (called SONHO), but running independent instances. This allows the Wrapper implementation to be reused at both sites. Each Hospital runs different Laboratory and Radiology information systems, which were integrated in RTS (from four different vendors). The primary care centres run independent instances of the SINUS information system as the main Patient management system (supplied by the public health sector administration agency). The present instalment covers a target community of ~1,000 physicians, from which ~35% have requested credentials for access.

The RTS deployment provides two portals for end-users: the Health Professionals portal (HP Portal) and the Citizen Portal.

6.1.2 THE HEALTH PROFESSIONALS PORTAL

The R-EHR is accessible using a web browser, without additional preparation from the client side. After authentication, the health professional queries for a given patient using common accession keys. Patients have a National Health Service number which provides a convenient key shared among partner institutions. The first view over the existing RTS clinical record is a list of episodes in a grid, providing essential facts such as time reference and data provenance (Figure 6.1): the patient is identified (bullet 1 in the illustration) and the known care episodes listed (bullet 4); an episode can cluster sub-episodes (bullet 2); the grid displays encounters from several institutions (bullet 3).

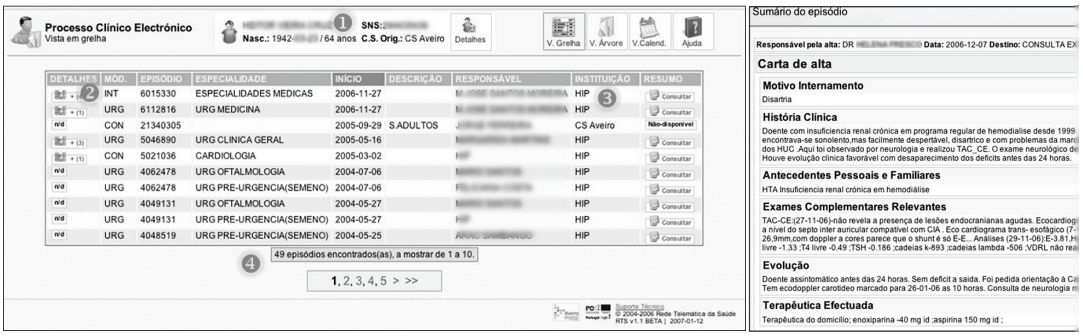


Figure 6.1: RTS Health Professionals' portal user interface (screenshot from the production environment; patient and physicians identifiers are blurred). On the right: a sample discharge letter normalized in RTS.

The user may then open each entry to access the details, for example, a discharge letter, a radiology report, or lab results. The expansion of an entry is typically a normalized summary, as specified by the working teams in the specification phase. Note that the summary information is prepared and displayed with the same look-and-feel as the rest of the portal. Instead of a summary, an entry may provide a direct link to the source information system (a URL to an external web). In this case, the accession keys of the patient and/or episode of interest are also encoded in the redirection link.

The portal is available in all the partner institutions, presenting the information retrieved from different information systems, scattered by different organizations in the RTS consortium.

6.1.3 THE MYRTS PORTAL

RTS offers a portal environment to the patient, the myRTS, which provides on-line services to registered users. Using myRTS, the patient is able to issue several requests to the RTS partner institutions, such as requesting consultation appointments and interact with the responsible General Practitioner. myRTS can be seen as a tool to help citizens manage their own health plans: it provides a calendar-style view of past and future events being automatically pooled from existing information systems in the partner institutions. The user can then expand each appointment and get additional information such as, for example, preparation instructions. Within myRTS, the citizen can also manage his/her vaccination record, using consolidated information from primary care partner institutions. In addition to the patient demographics and health agenda, myRTS allows the citizen to monitor the accesses being made to his own RTSys record; this option creates evidence to the patient of who accessed his record and when. The patient cannot upload health data elements, only 'read' selected event information, automatically gathered from the operational information systems.

Patient access is still in a pilot phase as the services provided in myRTS require that partner institutions implement the proper backend processes (*e.g.* designate pivot staff to answer requests made in the portal) In addition, the planned authentication strategy uses the National citizen identification card (http://www.cartaodecidadao.pt/), a smartcard that was not yet widely available during the deployment of RTS (nor the terminals to read it).

Table 6.2: Statistics on the current deployment of RTS (as for October 2011).

Connected source information systems: 8.
Users registered in the system (health professionals): 351. Active users: 53.
Accesses in October-2011 (sessions opened): 120.
Principal use of the system: GP access their patients' R-EHR to browse information produced in HIP.
Episodes available through the RTSys: ~19.000.000

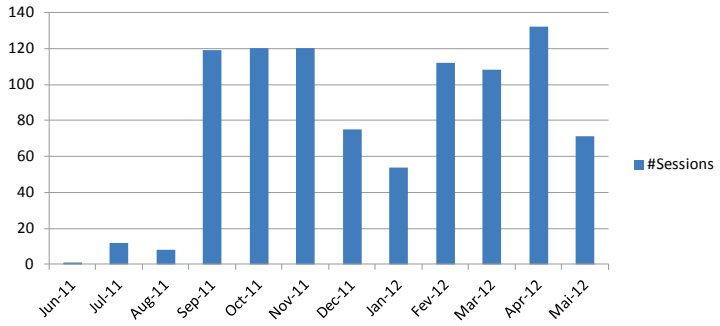


Figure 6.2: User sessions opened in the RTS (last year of operation).

6.2 Pilot use of RTS

The pilot use of RTS started early in 2007, with the HP Portal available in eight partner institutions. A large number of users (203) received specific training in the use of the portal, but it was not expected that many would use the system immediately, since there were two conditions yet to be met for the successful roll-out of the platform: (1) the Administrations should internally endorse the use of the platform and disseminate the associated practices among their professionals and (2) the information available in the R-EHR was to be extended to include critical patient summary characteristics (dependent from the collaboration of external entities). While these conditions were pending, the consortium decided to start the pilot use of RTS with the more enthusiastic users.

The users were keen to recognize the value of the platform and did praise its use, but those missing conditions constrained the use of the platform. In addition, a maintenance program was deemed necessary to ensure the safe use of the platform in an extended usage scenario, calling for action on the partner institutions. The practical conditions to implement the maintenance procedures were not available and the technical team at the University of Aveiro had to limit the RTSys development.

A few tens of users are using RTS in a regular basis. Though not fully explored, as initially intended, the tool is available and used in routine (Table 6.2, Figure 6.2), especially by the primary care professionals to browse clinical data originated in the Hospitals. The service level of the platform is now and then affected by external conditions (*e.g.* network reconfiguration) which, allied to an informal maintenance program, originates some periods of low usage.

6.3 Discussion

While the technologies and the standards exist to accomplish a longitudinal, multi-institution use of electronic patient records, strategic leadership, investments and political support are still lacking (Stroetmann *et al.*, 2009). The RTS was proposed in 2004 as a consortium between health care institutions in the region of Aveiro (Portugal) to implement a collaborative regional health network, linking Hospitals and primary care, sharing a common information space to support the continuity of care.

The inter-institutional approach in RTS, both with respect to the outcomes (community-wide online services) and the implementation methods (multi-disciplinary teams involving professionals from several organizations) was a novel model of work in the health care system, with respect to the specification and introduction of health information technology.

The requirements from RTS project led to the technical design of the RTSys in which the key application is a virtual, regional EHR, automatically built from patient data available at different systems at eight partner institutions. The R-EHR enables a coherent and comprehensive view over the fragmented patient data in the region. The proposed design is non-disruptive, preserving existing investments and responsibilities, while making it practical and secure to delivery shared EHR at the point of care.

At the core of the RTSys infrastructure is the HIETA middleware, supporting extensible and scalable integration of healthcare information systems. The middleware enables the abstraction of a common information model, without the need to change or adapt existing systems.

We have successfully developed a digital platform that enables the controlled and secure sharing of Patient data between care providers, for the continuity of care. This specialized 'system of systems' was a novel contribution in the National context, since no other regional health information network engine existed based on the same flexible and decoupling principles of RTSys. Project partners perceive RTS as the basis for building a broader health information infrastructure and as a demonstrator of the benefits this new approach brings to the Portuguese healthcare reality.

Besides the immediate application of RTSys to support point of care access to a broader patient record, the system has also been used as a research subject and a research tool. The work by Gomes *et al* illustrates the former, in which RTSys was the motivating context for the development of an authorization architecture for health applications (Gomes *et al.*, 2007). The later is best illustrated by the work of Ferreira *et al*, using natural language processing methods to analyse discharge letters, resulting in a thesis (Ferreira, 2011) and a book (Ferreira *et al.*, 2012).

The RTS has been awarded an innovation prize assigned by the industry and cited as an example of the advances in the Portuguese health information technology landscape, *e.g.* (Progress-Consulting *et al.*, 2011).

7 Bridging connected care with e-Science

In the previous chapters we have presented an infrastructure to enable connected care, in which different institutions agree to share clinical data in a secure and distributed telematic platform.

We will now address a different line of work directed to what can be freely called 'connected science'. Likewise, we are looking for a computing infrastructure to enable different people to form dynamic collaborations, in this case, in biomedical research activities for health applications.

We argue that the federated health information infrastructure implemented in RTSys and new advanced computing infrastructures for science can be used in complementary ways, enhancing the impact of RTSys as an information tool. The use of available science e-Infrastructures can enhance the care-oriented tools by providing, for example, the methods for research groups to aggregate data from multiple care sites (attaining critical mass for the study of less common diseases); new methods based on the use of demanding computing resources (*e.g.* simulation, 3D medical imaging rendering) can be delegated on powerful computing centres; and, perhaps more significant, the ability to integrate data from the clinical practice (structured in the EHR) with the emerging knowledge from biomedical databases (*e.g.* genetic bases of diseases). We believe that the clinical information systems and the (yet) research-oriented biomedical resources should be used in complementary ways and this is an open challenge to biomedical informatics.

7.1 Concepts and opportunities in e-Science and Grid computing

The term e-Science denotes the application of research methods based on the use of advanced computing resources (Hey *et al.*, 2003). These resources may include computing capacities (processing), very large data sets, expensive scientific instruments (*e.g.* medical imaging scanners), and advanced visualization. Perhaps the most striking example is the use of information technology and computer methods in the Large Hadron Collider experiments at CERN (Bird *et al.*, 2009). The scale of data production, transfer and analysis in this context is something without precedent, requiring new computer methods to enable breakthroughs in (physics) science (Geddes, 2012). The crucial role of ICT to manage the 'data deluge' in the major scientific challenges ahead is well pictured in the literature (The 2020 Science Group, 2005; Hey *et al.*, 2009).

A specific technology to support scientific computing is the Grid. Grid computing has been proposed to enable distributed computing infrastructures, allowing multi-institution communities to share high-end resources and practices (Foster *et al.*, 2001). The Grid provides a viable approach to deal with the data deluge found in modern science in many domains (Gray *et al.*, 2005), such as life sciences or environment studies (illustrative application examples are available from the European Grid Infrastructure web site: http://www.egi.eu/case-studies/).

The use of Grids for scientific computing aim at facilitating: (1) the execution of comprehensive experiments, using large datasets and complex algorithms; (2) the collaboration between different research centres to attain higher critical mass; (3) structuring thematic communities, allowing them to share best practices; (4) optimize the use of expensive computing resources and scientific instruments.

Any organization can set up a Grid infrastructure, but the natural deployment of Grids has a much broader scope, aggregating advanced computing centres (usually benefiting from public funding) under a governance model that favours fair access by researchers (Bird *et al.*, 2009). That is why we find Grid deployments usually related to national research authorities (*e.g.* Open Science Grid (http://www.opensciencegrid.org/), in USA; NAREGI (http://www.naregi.org/index_e.html), in Japan; eScience (http://www.nesc.ac.uk/), in United Kingdom)). Portugal shares efforts and resources with Spain in a regional initiative, forming the IberGrid infrastructure (http://ibergrid.lip.pt). Our research group has been operating a local node in the IberGrid community, since its set up.

The European Commission has strongly supported the adoption of Grid computing as an instrument to cluster scientific communities in the European Research Area, under the e-Infrastructures initiative (see, for example, Annex A of the SIENA roadmap (SIENA, 2012)). Currently, almost all the national Grid infrastructures in Europe are federated under the umbrella of the European Grid Infrastructure (EGI, http://www.egi.eu), including the IberGrid, supporting communities in domains such as biomedicine, physics, astronomy and chemistry (http://www.egi.eu/community/vrcs/).

The Grid architecture: software services to access distributed capabilities

Grids are enabled by a specific distributed computing middleware that provides the required abstractions to isolate applications from resources provision (Figure 7.1), making available high-level access services to submit computational jobs, manage data (input and results), and track the operational status of the infrastructure. The typical components of the middleware are depicted in the central layer of Figure 7.1 (adapted from the gLite middleware, http://glite.cern.ch/). An authoritative discussion about the architectural principles for the design of Grid middleware is available from Foster (Foster *et al.*, 2003).

Leading Grid middlewares include the Globus Toolkit (http://www.globus.org/toolkit/), a pioneer effort, and the European Middleware Initiative (EMI, http://www.eu-emi.eu/) that resulted from the research required to develop the computing methods to support the LHC experiments (Geddes, 2012). The EMI distribution is an evolution of gLite, now being discontinued.

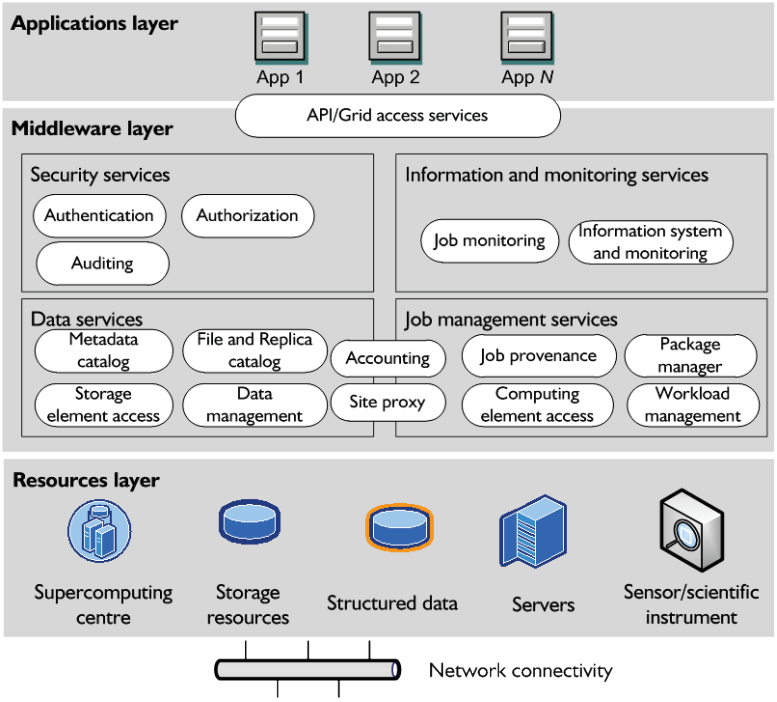


Figure 7.1: Multi-layer architecture of the Grid infrastructures. The middleware makes heterogeneous high-end resources usable to applications in a controlled environment.

The need to reconcile the emerging Grid middlewares and infrastructures led to the formation of the Open Grid Forum (http://www.gridforum.org/) in charge of consolidating the community best practices into standards. One of the most relevant technical standards is the Open Grid Services Architecture (OGSA), defining the middleware as a three layer distributed systems model (access layer, services layer and resources layer), as illustrated by Figure 7.1, and the scope of the capabilities required to enable Grid systems and applications (OGF).

Virtual organizations: collaboration spaces sharing similar requirements

While other distributed computing infrastructures exist, Grid addresses the specific problem of 'coordinated resource sharing and problem solving in dynamic, multi-institutional virtual organizations' (Foster *et al.*, 2001). The Virtual Organization (VO), a central concept to the Grid computing, is a dynamic association of people or institutions with similar scientific interests and requirements, and provides a key abstraction to structure the controlled use of Grid infrastructures, both from social and technological perspectives. Our group, for example, participates in the large biomedical VO in the EGI, and in the smaller Brain Imaging Network Grid VO in the IberGrid.

The composition of a VO is managed by specific membership services in the Grid middleware. Users of the Grid infrastructure must first be a member accredited in one or more supported VOs and authenticate themselves by providing strong digital credentials, by using X.509 certificates (Housley *et al.*, 2002).

Grid and Cloud

The current buzzword for the deployment of scalable computing infrastructures is Cloud computing (Armbrust *et al.*, 2010), but, in fact, the concept is not new and share many ideas already found in the well established Grid computing concept (Foster *et al.*, 2008).

Grid and Cloud have a lot in common with respect to the general vision, architecture and enabling technologies. Both are based on the use of standard web technology and appropriate middleware to provide on-demand a pool of abstracted computing and storage power to consumers.

There are differences too, starting with the business model. Clouds are more based on central infrastructures, controlled by a large operator that sells computing as an utility (processing, storage and communications) to any consumer. The Grid is specially targeting the federation of high-end scientific infrastructures, under decentralize control. This very last issue also justifies different security models in both paradigms. Concerning the programming models, the Cloud uses mainly a services approach (based on Web Services (Alonso *et al.*, 2004)) and scripting technologies; the Grid supports more directly the use of high-performance scientific computations paradigms, such as the explicit use of parallel computing (Foster *et al.*, 2008).

Clouds are rapidly ramping up as a deployment infrastructure of choice of many modern internet-aware applications. Clouds can, naturally, be used as a resource for scientific computing, and, in this sense, provide an alternative to the use of Grid computing (SIENA, 2012; Rosenthal *et al.*, 2010). But there is also space for synergies, perhaps the most obvious is the use of Clouds for the provision of the infrastructure to deploy Grids; this is the object, for example, of European project StratusLab (http://www.stratuslab.org/).

The use of Clouds in scientific applications is a new field (a partnership for the development of the 'science cloud', involving three leading research organizations in Europe, has been announced early in 2012 (HN-SC)). A promising direction is the use of federation models for research activities involving Cloud infrastructures (*e.g.* EGI-inSPIRE, http://www.egi.eu/about/egi-inspire/), a successful paradigm that has been used to aggregate Grid sites for e-Science.

The rapid evolution in this area will introduce new opportunities to rethink the models we explore in this chapter (for the Grid paradigm), not specifically concerning their migration into the Cloud, but with respect to the future technologies to come, more likely to be an evolution of the Grid and Cloud paradigms.

HealthGrid: Grids for health applications and biomedical research

The use of Grid technology in the life sciences domain for health applications and biomedical research is labelled as 'HealthGrid' (Breton *et al.*, 2005). The full realization of the HealthGrid vision, as promoted by the European Research Area, is a cooperative yet controlled, distributed yet secure, e-Infrastructure capable of access and relate life science data at multiple levels (Andoulsi *et al.*, 2008). The European Commission proposed a five domains model for the integration of information at different scales and disciplines: molecule, cell, organ, individual and

population (Figure 7.2). Such vision implies the practical use of a large number of digital information sources in a resourceful computing environment, ultimately contributing to a comprehensive understanding of the mechanisms conditioning life (and health). The HealthGrid is seen as a promising instrument towards a holistic use of information about patients and pathologies, exploring the cross-fertilization between disciplines (Figure 7.2) (Breton *et al.*, 2005; Andoulsi *et al.*, 2008).

There are several examples of successful HealthGrids, including modelling the human body for surgery planning, pharmaceutical research and medical image processing. Refer to the SHARE Roadmap for a comprehensive review (Andoulsi *et al.*, 2008) and to the work of Sridhar *et al* for a critical appraisal (Gwadry-Sridhar *et al.*, 2010)).

Medical imaging process is one of the most relevant Grid applications for health (Olabarriaga *et al.*, 2010) and has also been pursued in our group, addressing the requirements of brain imaging (Cunha *et al.*, 2007), endoscopic capsule exams processing (Oliveira *et al.*, 2010) and medical image repositories (Oliveira *et al.*, 2008).

7.2 Complementarities between care and science infrastructures

The secondary use of health data originated in clinical practice is essential for medical knowledge development (Safran *et al.*, 2007). This means using the clinical data to feed research workflows, for example, providing clinical evidence from the aggregate use of a large extent of health records. Research data is usually not required to be identified, facilitating the secondary use of the EHR content.

On the other hand, advances in biomedical knowledge, specially the evolution of the understanding of molecular processes and genomics, are valuable to enrich the clinical practice, but not easily embeddable into current information tools (Martin-Sanchez *et al.*, 2004).

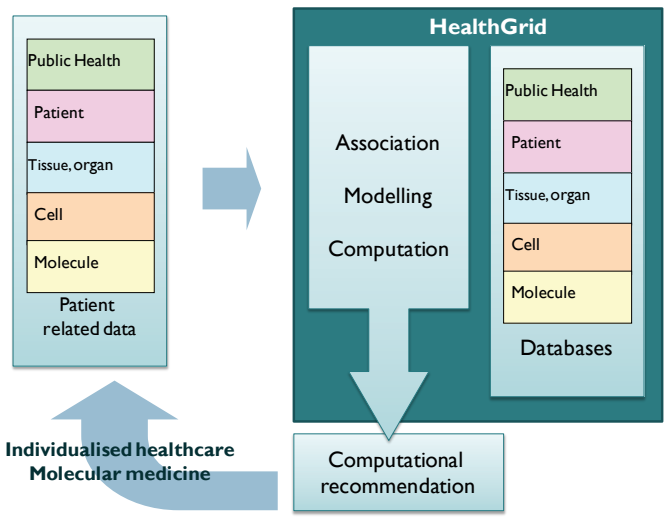


Figure 7.2: A conceptual roadmap for HealthGrids research: the full spectrum of biomedical information supports personalized care (credit: S. Norager, Y. Paindaveine, European Comission).

Several authors have noticed that the two domains (health tools for clinical practice and biomedical knowledge captured in databases) should be integrated in practical ways (Kohane, 2000; Martin-Sanchez *et al.*, 2004). This was the central research goal of the Infogenmed project, in which our group has previously participated (Pereira *et al.*, 2002).

In the following sections we study how the RTSys (a health domain tool) could be linked to the global biomedical knowledge, by leveraging on existing infrastructures for e-Science.

7.2.1 INTEGRATION OF BIOMEDICAL INFORMATION RESOURCES FOR HEALTH APPLICATIONS[23]

The recent advances in genomics and proteomics raise an enormous potential to change clinical practice in which diagnosis and treatments will be supported by information at molecular level, towards personalized assessment and therapeutics (Collins *et al.*, 2001; Martin-Sanchez *et al.*, 2004). This raises new challenges to the information tools, traditionally developed apart for the medical and biological fields, requiring an increased exchange of knowledge between the two domains (Martin-Sanchez *et al.*, 2004; Altman, 1998). While the former has a deep knowledge of the *phenotypes*, the last has the practice and the tools to analyze the *genotype*.

Linking the phenotype and genotype requires new information tools for seamless distributed biomedical data integration, since the existing ones are not ready to link genetic and clinical information in support of the clinical workflows (Mitchell *et al.*, 2003). This means that patient data, information about pathologies, clinical trials, genetic sequences, proteins, etc, should be seamlessly navigable to practitioners, based on adequate computer methods.

Requirements for biomedical information integration

Information integration refers to the problem of merging disparate information sources such that they appear to a user as a single coherent source (Gupta *et al.*, 2003). It is not a new problem and has been addressed previously by a vast number of systems (Parent *et al.*, 1998). The integration of modern biomedical knowledge, however, brings up a new problem domain with some specific challenges:

- There are many different sources of information spread over the Web; the relevant information needs to be located, accessed, and retrieved.
- Data integration is difficult since databases can present a wide range of formats and different semantics. In addition, public information resources are often only available through web interfaces, not easily interrogated by computer methods.
- Coding and terminologies are not unified, sometimes being difficult to discern quality and link related concepts. In spite of the normalization initiatives, gene naming, for example, is far from being standard.

[23] Partially adapted from our previous work published in Oliveira, I. C., Oliveira, J. L., Sanchez, J. P., López-Alonso, V., Martin-Sanchez, F., Maojo, V. & Sousa Pereira, A. (2005). Grid requirements for the integration of biomedical information resources for health applications. *Methods of Information in Medicine*, 44(2), pp. 161-167.

— Medical coding systems are not ready for managing the emerging genetic information.

— Existing bioinformatics tools are mainly designed for researchers, not for health practitioners.

— Physicians often lack guidance to select and access the relevant data from web sites that are very specific in areas they are unfamiliar with or geared to biological experts (Mitchell *et al.*, 2003).

— Intellectual property rights, privacy and confidentiality issues and protection of the ownership of valuable data may hinder the exchange of contents.

— Results are often published in natural language formats (scientific bibliography), requiring mining techniques to recover the knowledge in computer ready representations.

Among the problems identified, the most fundamental technical requirement to address is the semantic integration of information from multiple domains and layers; the problem is highly complex for the domain of medical informatics alone (Stroetmann *et al.*, 2009) and gets more complicated with the inclusion of the bioinformatics domain, though relevant semantic tools exist in both fields (*e.g.* the GeneOntology (Ashburner *et al.*, 2000) provides the 'semantic glue' to link different biomedical concepts).

Integrative computing infrastructure requirements

From the lessons learned in our previous work in the Infogenmed project, the most pressing requirements for a distributed ICT infrastructure to connect the health applications and new biomedical knowledge resources are those related to distributed information gathering, transformation and unification, including:

— Advanced data access abstractions, enabling a standard service-interface over heterogeneous information sources, in spite of the variability in the capabilities of the underlying data management systems. Such interface would enable the access using a standard query language over a virtual information model (Wöhrer *et al.*, 2005).

— Secure transport and, in some cases, secure storage of sensitive data. It is worth to stress the absolute need to observe privacy, confidentiality, legal issues and intellectual property provisions (Andoulsi *et al.*, 2008).

— Standard languages for sources description, enabling to produce and share semantic-rich models. The information sources, as they became available, would then register as content providers, describing the information they hold according to the shared semantics. To this end, the definition and use of shared ontologies and terminology systems and the reuse of existent ones (especially in medical informatics) are essential. These requirements are in line with the 'semantic Grid' concept (De Roure *et al.*, 2005).

— Decentralized and dynamic source registration and new content discovery.

— Distributed query planning and execution will become an essential component in knowledge integration systems in life sciences. Moreover, advanced distributed data

services (Venugopal *et al.*, 2006), such as subscriptions and notifications, content expiration, data provenance, etc., are valuable.

— Flexible sharing models, enabling fine-grain access policies. One should note that the integration of medical and genetic data encompasses both private and public databases and, in some cases, paid content.

These requirements have been discussed by the Grid community and several complementary analysis are available in the literature (Hastings *et al.*, 2002; Breton *et al.*, 2005).

7.2.2 APPLYING GRID TECHNOLOGY IN BIOMEDICAL INFORMATION INTEGRATION

The integration of the existing biomedical information for medical applications requires establishing semantic connections between highly heterogeneous resources, from different disciplines, at the global scale (Martin-Sanchez *et al.*, 2004). The Semantic Web (Berners-Lee *et al.*, 2001) technologies can help with the semantic description of resources to fuse on-line information (Lopes *et al.*, 2011), but they fail to address the specific issues of infrastructure deployment and management of 'trusted federations' (of research centres).

The Grid, enhanced with semantic integration services, is expected to contribute to the integration of data sources and knowledge between the clinical practice and the genomic research (Andoulsi *et al.*, 2008; Tsiknakis *et al.*, 2008; Montagnat *et al.*, 2008a). The Grid technology already supports (1) the provision of pervasive computational power on-demand, (2) strategies to manage unstructured (file-oriented) and structured data (especially relational), and (3) the management of collaboration spaces of federated resources and communities (Virtual Organizations).

To harness from the Grid for the integration of biomedical resources, including bridging from the clinical practice to the "omics" world, the basic infrastructure must be complemented with semantic services, enabling the use of shared representations and the semantic linking between concepts originating in different sources (Cannataro *et al.*, 2005). This can be achieved by extending the Grid to include domain ontologies, semantic annotation of resources and services, and semantic reasoning tools, towards a 'Semantic Grid' (De Roure *et al.*, 2005). There are several projects exploring this line of work. The Cancer Biomedical Informatics Grid (caBIG) is a collaborative Grid-enabled network aiming at fostering the collaboration between cancer researchers, enabling the community to exchange data and knowledge (Oster *et al.*, 2008). The Biomedical Informatics Research Network (BIRN) infrastructure uses the Grid to support the neuroscience research community (Helmer *et al.*, 2011). It provides the means for researchers to federate and analyze data from different neuroimaging sites, making intensive use of medical imaging modalities and genomic data. The link to the clinical practice is perhaps best illustrated by the @neurist project, proposing the use of the Grid to integrate clinical and genetic data, conduct simulations and support clinical risk assessment, in the domain of intracranial aneurysms (Benkner *et al.*, 2010). Another good example of the advanced used of Grid as an information integration platform is the Health-e-Child project, proposing a Grid-enabled system to integrate large-scale, multi-modality data to support medical decision support in the field of paediatrics (Freund *et al.*, 2006).

7.2.3 AN ARCHITECTURE TO EXTEND RTSYS TO THE GRID

RTSys is a health-information network platform to support the clinical practice. Its value proposition builds on the integrated access to patient information in a community, adding value to the information (which is more comprehensive) for clinical decision. We can identify two possible vectors for the enhancement of RTSys that intersect with the value proposition of research e-Infrastructures:

1) Enrich the functional scope of the health information network by using RTSys as a familiar entry point to access advanced computing infrastructures. Candidate use cases in this kind of integration would include (i) resorting to advanced simulations or rich visualizations, not easily attainable with local computational resources or local software tools; (ii) global data integration, including new sources of biomedical knowledge, often publicly available in the Internet; or (iii) data mining methods for knowledge extraction. The regular user, departing from a specific patient context (the R-EHR s/he is browsing) would be able, for example, to build a 'disease mind-map' (Dias *et al.*, 2006) on that subject, by having a remote, robust platform collecting and integrating information from multiple databases to deliver a risk assessment index. Or, more prosaically, have a given modality remote analysed for automated detection, for example. This integration could provide the practical support to navigate from the patient's actual phenotype to the genotype, by semantically relate the R-EHR with remote global biomedical resources.

2) Open the data accessible in RTSys to research communities (as long as the appropriate data protection mechanisms are in place). Although the medical imaging modalities are not being shared in the current RTSys deployment, they can be included in the sharing practices and provide, for example, valuable raw data to a bank of clinical cases for secondary use.

These scenarios correspond to a vision in which RTSys would integrate (acting as a client) with an advanced research infrastructure offering the characteristics of a HealthGrid. From a system architecture point of view, there are new developments required to interface RTSys and the Grid, especially to enable data flow between the clinical domain and the e-Science domain, including the ability to protect the privacy of the medical data leaving the platform (Figure 7.3). Since the expected use cases would likely include clinical decision support scenarios, it would be important to support re-linking the results back to the patient. In this kind of application, the data fragments from the R-EHR should be pseudonymized, storing locally the mapping to re-establish the patient identity (Elger *et al.*, 2010).

Another new module in RTSys should address the required 'interoperability enhancement', *i.e.* the ability to export the R-EHR in different standards, according to the application scenario and the clinical representations available from the Grid components (Figure 7.3). This would involve some fundamental choices concerning the representations to be shared by the computational methods. It would be more realistic to expect that the data leaving the RTSys should be aligned with a more encompassing domain model or ontolology defined for the Grid-enabled tools. It could be necessary, for example, to transcode parts of the R-EHR to a new content model. The use of medical images is of special interest given the high requirements

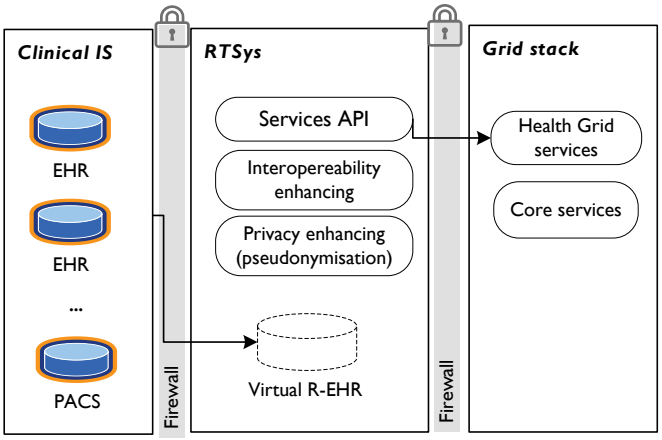


Figure 7.3: The RTSys exposes data accessible to the Grid middleware.

involved in storage and processing and the availability of standard interfaces to access them (which also call for specific approaches for data management and transfer). Several groups have addressed the use of medical images in the Grid that are available at clinical settings (Montagnat *et al.*, 2008a; Blanquer *et al.*, 2010).

The existing services API of RTSys can provide the required connectivity to expose regional data to remote infrastructures. The information structures included in the services available should be extended to address the previous requirement.

From the Grid stack point of view, there are also requirements for new features to equip the standard Grid middleware up to the level of a semantic HealthGrid, able to relate global biomedical sources of knowledge (Figure 7.4).

The bottom layer (Figure 7.4) corresponds to the computation resources distributed infrastructure, including the hardware capacities and connectivity to ensure high-end processing and robust data storage. (The resources layer can conceptually include high-end medical instruments but, for this discussion, we look into this layer as a black-box of computing resources).

On top of this, an interoperable Grid middleware, conforming to the services model defined by the Open-Grid Services Architecture (OGSA)(OGF) ensures the key Grid functions: execution and workload management, (domain-independent) data management, infrastructure information system and the Virtual Organization management. This middleware can be provided by the Globus Toolkit (http://www.globus.org/toolkit/), the European Middleware Initiative (http://www.eu-emi.eu/) or another stable and mature Grid middleware.

The core Grid functions need to be supplemented by domain specific components. These are extensions components integrated with the core to provide biomedical data representations and algorithms (Figure 7.4). In this services layer there are components that extend the normal functions of the workload, data and security management systems.

The proposed components, following a SOA approach (Papazoglou *et al.*, 2007), are the following:

Medical data processing. A repository of operators on medical data, described in reference to the common semantic models, is available to use on analysis workflows. These are Grid-enabled algorithms, meaning they are designed to use the distributed high-throughput nature of the infrastructure. A group of operators frequently proposed is formed by computer methods to process imaging modalities (Montagnat *et al.*, 2008b).

Workflow and high-level scheduling. Biomedical analysis often implies specific stepwise protocols, not addressed by a batch approach commonly found on current Grids. The services enabling clinical workflows build on the core execution manager and add better application-level and data-aware scheduling policies (Deelman *et al.*, 2009).

Ontologies and Semantic models. The use of ontologies enables the explicit modelling of semantic relationships between domain concepts, required to ensure that computer methods can automatically relate content from multiple resources. The source schemas and the processing services should them be described (annotated) to ensure the concepts they manage can be linked to the shared representation in the ontology, enabling for services discovery, rich query interfaces and knowledge extraction tools (Cannataro *et al.*, 2004).

This component plays a critical role in the solution, since it provides the means to link knowledge produced in different sources, from different disciplines and includes three different parts: an ontology, the annotation of Grid services and the annotation of data sources' schemas. Since it is often very complex to find a widely accepted and comprehensive ontology, the solution may resort to an integrative ontology, building on existing ones or, as a shortcut, the replacement of the ontological framework by a reference domain model.

The semantic models enable the formulation of queries at the concept level (expressed in the terms of the ontology) and their mapping to the specific data access scenarios (*e.g.* translating into Structured Query Language understandable by a given source). The data annotation is essential to enable a semantic mediation between heterogeneous sources.

Biomedical sources access. The access to clinical data modalities may need specific methods

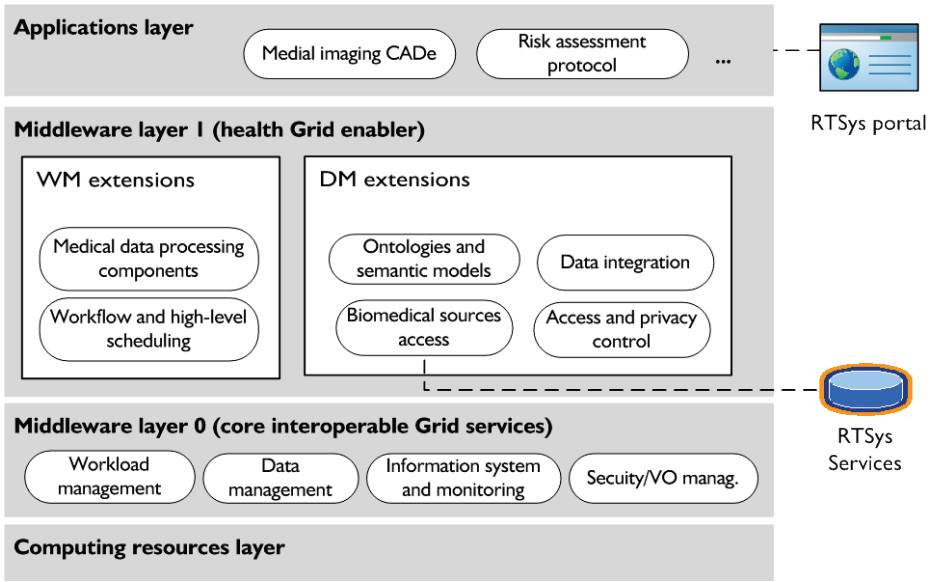


Figure 7.4: Service layers to enable a biomedical Grid.

to handle the technical interfacing and syntax and expose the data through services and information models convenient for the data integration procedures. This component would include a collection of specialized 'data wrappers', following a similar approach to the one adopted in RTSys.

Data integration. The role of data mediation services is to offer the abstraction of a virtual, coherent information model by integrating remote sources. The source systems fully preserve their autonomy and it is up to the data mediation services to address the technical and semantic heterogeneities. To this end, the mediation is expected to heavily rely on the *de facto* standard Grid services for structured distributed data access and query processing provided by Open Grid Services Architecture Data Access and Integration/Distributed Query Processing suite (OGSA-DAI/DQP) (Alpdemir *et al.*, 2005), although its support is dependent on the middleware choice.

Access and Privacy control. The generic VO concept provides a community-level view for authentication and authorization, centred on the access to computational resources. The access and analysis of medical data implies a role-based access control (Doosselaere *et al.*, 2008). In addition, the use of clinical data in a massively distributed infrastructure should apply privacy enhancement techniques, hiding the identity of subjects (and not blocking the flow of data). Pseudonymization provides a way to remove direct identifiers from medical modalities (medical images, clinical reports, etc.), as they leave the clinical domain, with the possibility to link the results back to the clinical practice in a later stage (Elger *et al.*, 2010). Privacy protection of clinical data is not trivial; while there are obvious identifiers that can be swapped, the content of data can be itself a source of identification (*e.g.* high-resolution MRI scans allow to reconstruct a person's face (Toga, 2002)). New privacy enhancing and fine-grain access control services should be introduced to complement the basic Grid security components (Kalra *et al.*, 2005).

The top layer (Figure 7.4) addresses the application specific tools to implement selected research or clinical use cases. The applications listed are just examples and not concrete proposals: computer aided detection of features on medical imaging using the Grid has proved its value (Bellotti *et al.*, 2007); the integrative application of the Grid for global data access, linking private Clinical IS and public biomedical is a comprehensive example of the possible applications (Benkner *et al.*, 2010).

The layered approach to structure the Grid stack (Figure 7.4), including a domain-specific services layer for the biomedical domain, can be found in the work of other groups, as in the architecture of the BIRN landmark collaboration in neuroscience (Helmer *et al.*, 2011) and the knowledge-enabled Grid serving the @neurist project (Benkner *et al.*, 2010). Both projects resort to the OGSA-DAI/DQP approach to implement the required data integration of global data sources.

7.3 Exploratory studies and prototypes

Connecting RTSys and the e-Science infrastructures is an ambitious vision, imposing a very demanding set of developments in the Grid side to attain the level of sophistication expected from

a biomedical knowledge Grid. Nevertheless, we have explored different opportunities to take steps on that direction, working on the theoretical architecture and partial validation experiments, reported in the following sections.

7.3.1 CONNECTING THE CLINICAL PRACTICE TO THE GRID IN ENDOSCOPIC VIDEO ANALYSIS

As a proof of concept, we conducted an experimental prototype to Grid-enable a clinical tool used by medical doctors in routine. The overall challenge was to move the execution of a demanding detection algorithm used in a desktop endoscopic video analysis tool and have it run on a remote Grid infrastructure. The methods and results are reported elsewhere (Oliveira *et al.*, 2011) and we will provide a brief presentation of the prototype here.

The target application scenario

The Wireless Endoscopic Capsule (WEC) is a medical device used to record a video of the inner body (Qureshi, 2004). The capsule is swallowed and produces a wireless video signal as it travels though the individual's inner bowel, captured by a wearable recording device worn at the waist area. It has the practical drawback of producing a video 6 to 8 hours long, which is analysed in a desktop computer, taking about 2 hours to fully review.

Previous research in our group led to the development of the CapView software (http://www.capview.org/), a desktop application to browse the WEC video. CapView adds some convenient features to assist the clinician, such as the ability to automatically establish the location of the four main topographic regions (entrance, stomach, small intestine, large intestine) and also calculate the capsule transit times, applying the Automated Topographic Segmentation (ATS) algorithm (Cunha *et al.*, 2008) . The ATS algorithm is based on computer vision methods to find the four main anatomic sections of interest, using Support Vector Machine (SVM) classification (Burges, 1998) applied to MPEG-7 descriptor vectors (Chang *et al.*, 2001); refer to (Cunha *et al.*, 2008) for details. The single image classification stage iterates over each frame, for an average count of 60,000 frames per exam, and classifies it based on previous trained model using MPEG-7 scalable colour features. After the classification of each frame, a segmentation stage applies a global model fitting approach to estimate the position of the oesophagus-gastric junction, the pylorus and the ileo-cecal valve based on zone transitions. The locations are then written in a XML file and are ready to open in the *CapView* software. This detection process takes about one hour to complete on a standard desktop computer (for an average size exam, *i.e.* ~500MB).

Grid integration experiment

The video to be analyzed for each exam is encoded in Motion-JPEG (which natively offers random-access to any frame) and the single image classification of the ATS algorithm analyses independently each video frame (to discriminate the zone that frame belongs to). Given these premises, it is both feasible and practical to split the original video in segments and analyze them concurrently. A simple 'bag of tasks' approach with data partition is used, since the analysis of

each video frame is independent. The original dataset (one large video) is divided into smaller segments and distributed over several Grid working nodes, running the same operator (single image classification) in parallel. During the processing stage, there is no need for communication between tasks due to the nature of the algorithm. The steps involved in the submission and remote invocation are the following:

1) Using the CapView familiar software, the clinical user orders the anatomic segmentation (selecting the corresponding option from the application menu).
2) The video is transferred to a demilitarized staging area (according to the predefined configuration settings), supplemented with a manifest file with metadata and the intended partitioning strategy. No direct access to the Grid infrastructure is required and no personal identification is sent (only local identifiers of the exam).
3) A helper application detects if there are unprocessed exams on the staging area; if so, it uses the settings in the manifest to partition the video and uploads the corresponding segments to the Grid storage element and prepares each job submission.
4) The jobs are submitted to the Grid workload manager to be run in parallel. The number of jobs is dependent on the partition strategy.
5) Each node runs the single image classification task on each frame of its assigned video segment.
6) The progress of jobs is monitored by the helper application (using the Grid information system) and, when a job is successfully completed, the output results are transferred to the shared staging area.
7) The CapView application picks the new results as they become available at the staging area and provides visual feedback to the user, marking the four anatomic parts (Figure 7.5).

In this experiment, we use a Grid interfacing library developed in our group, the IEETA Grid Framework (IGF), which allows invoking the gLite middleware, exposing a developer-friendly Java API. The IGF and this experience have been executed over the European Grid Infrastructure (EGI).

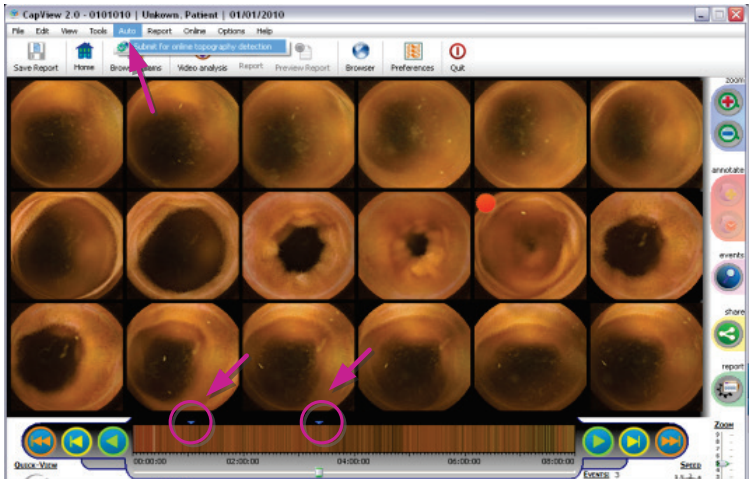


Figure 7.5: The CapView user interface. The menu option 'submit to remote analysis' allows sending the video to a Grid infrastructure for automatic topographic segmentation. Afterwards, the results (detections) are integrated in the visualization (blue dots, in the bottom).

Prototype results

Grid infrastructures for science are available and provide extensive amounts of resources but require expertise on its architecture and operations to develop even a simple application. Our experiment demonstrates the use of a familiar desktop application (CapView) to interface with the Grid infrastructure (Figure 7.5). The domain user does not have to possess any knowledge on the Grid technologies.

The results show that the large exams can be processed entirely with an average time of around eleven minutes, which is considerably faster than the desktop-based run. One should note that the overall algorithm does not scale well, as it requires merging all partial results before the final segmentation step. This makes the completion of the process dependent on outliers.

The use of Grid is appealing for research or clinical workflows involving a large collection of cases or repeated runs (*e.g.* with different parameters). In the case of the WEC video, a data partition fits the Grid programming model and can be used for massive parallel processing with clear advantages for the clinical end user or the researcher.

7.3.2 GRID-ENABLED IMAGING REPOSITORIES FOR MEDICAL APPLICATIONS[24]

Pursuing the line of research previously introduced, concerning the use in research of medical modalities from the clinical practice, we have worked on Grid-enabled models to deploy scientific repositories suitable for the cardiology, gastroenterology and neuroscience domains, exploring the commonalities related to medical imaging processing. This research was conducted in the scope of the GERES-med project[25], funded by the *Fundação para a Ciência e Tecnologia*. In this context, we worked with the specific domain communities to identify the most relevant use cases that should be supported by a 'repository of cases' infrastructure for e-Science, enabling the long term-storage and access in research activities.

The main contribution to this vision, in which we have worked, is the functional description of the Grid-enabled solution and the envisaged system architecture, described in (Oliveira *et al.*, 2008).

Use cases in medical repositories for e-Science

Medical images and text reports are an essential resource for modern diagnosis in several medical areas. Specialized clinical centres can easily reach the terabyte scale of patient data. While these information sources hold an amazing potential for training, education and research, they are often oriented to particular patient care and locked inside each Centre. Two good examples are the cardiology and the gastroenterology medical communities, producing and managing several modalities, such as dynamic images (*e.g.* videos) which requires high-

[24] This section is partially adapted from our work published in Oliveira, I. C., Fernandes, J. M., Alves, L., Pereira, A. S. & Cunha, J. P. S. (2008). GERES-med: An Architecture for Grid-Enabled scientific REpositorieS for medical applications. *In:* Silva, F., Barreira, G. & Ribeiro, L., eds. Ibergrid: 2nd Iberian Grid Infrastructure Conference Proceedings, 2008. Netbiblo, 163-173.

[25] http://www.ieeta.pt/sias/projects_Details.php?id=24

performance tools, with challenging problems concerning the acquisition, archival, processing and transmission (Figure 7.6).

The setup of clinical case repositories can benefit the clinical practice, but specially research and education. The following motivating scenarios were identified from the perspective of the cardiology and gastroenterology communities:

— The clinical expert located at the specialized clinical centre tag patient studies to share, published in the repository as pseudonymized cases. This includes medical images (*e.g.* CT, Echocardiograms, angiographies, endoscopies in MPEG-4 videos), reports on those images and clinical information and demographics.

— Doctors query the repository to find cases by similarity (*e.g.* with similar image features, with similar clinical history) to attain additional evidence for decision support. Ideally, computer methods would support content-based querying.

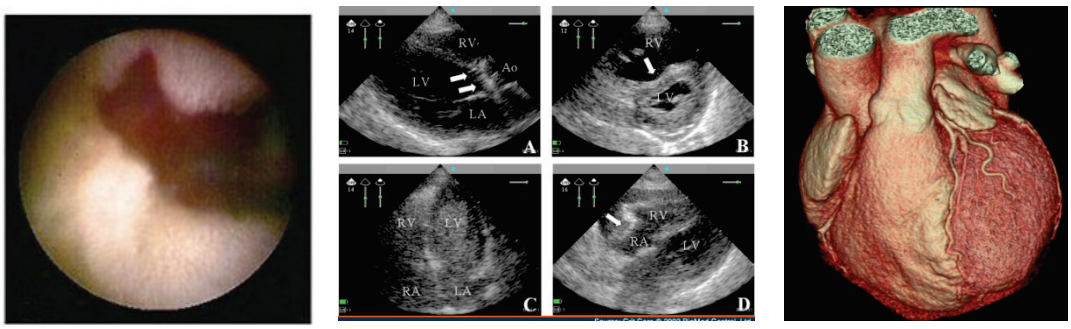
— The academic user navigates among classified cases in order to learn and acquire clinical practice.

— The researcher looks for training and test data sets to develop and test new image processing methods (that may be integrated as a new contribution in the framework).

Supporting these use cases over high-quality, long-term clinical repositories requires an advanced modelling of the semantics of data and tools. The use of semantic models and metadata plays a central role in distributed data sources integration and the ability to offer concept-level querying interfaces.

A medical research repository architecture proposal

The Grid computing paradigm provides technical solutions to address the federation of data in a dependable infrastructure. Grids are used for the (1) integration of demanding medical modalities over distributed data sources (e.g. (Saltz *et al.*, 2006; Keator *et al.*, 2008)), maintaining the autonomy of existing domains; (2) to run complex analysis methods, including Content Based Image Retrieval (CBIR) mechanism (Montagnat *et al.*, 2005); and (3) to semantically combine information at multiple levels of the biomedical knowledge landscape (Tsiknakis *et al.*, 2008).



A) B) C)

Figure 7.6: Wireless endoscopic capsule is used to produce videos of the inner bowel and detect problems, such as bleedings (A). Cardiac imaging requires high-performance tools especially for video based modalities such as Echocardiography (A) and 64-Slice CT scans (C, from Upenn Medicine[1]).

The Grid infrastructures are ready to offer standards-based processing services and reliable distributed storage for generic scientific applications, but the health domain brings specificities and additional requirements (Andoulsi *et al.*, 2008). Advanced data management is central to this purpose. There is a significant work in the area of medical Grid-enabled data management exploring methods of storage, replica management, metadata treatment and secure access (*e.g.* (Montagnat *et al.*, 2008a; Warren *et al.*, 2007; Erberich *et al.*, 2007; Blanquer Espert *et al.*, 2009)). The OGSA-DAI project allows structured data to be queried, updated, transformed and delivered using web services that can be deployed in Grid environments. This is an extensible and open framework used in many Grid-enabled projects (*e.g.* eDiamond, caGrid, BIRN), including the access to medical records (Liu *et al.*, 2010).

System architecture

We have presented previously a system architecture to enable the use of the Grid for health applications (section 7.2.3). The generic architecture then discussed can effectively support the needs of a case-oriented repository deployment, but includes some demanding components hard to achieve (*e.g.* the comprehensive ontology). For the present discussion, we analyse a simplified version of the global model presented in Figure 7.4 and highlight other specific functions.

A Grid-based solution to the long term deployment of scientific repositories of medical cases would include four main layers: (1) a web portal for the access layer, (2) a SOA interface, (3) Grid extensions for the medical repository management, and the (4) Grid Infrastructure (Figure 7.7). These components are presented below.

1. **Web portal**. All the resources and workflows are accessible through a user-friendly web portal. This portal will be focused on the high level use cases identified previously and provide the end-user with research-friendly abstractions ('my experiments', 'my cases',

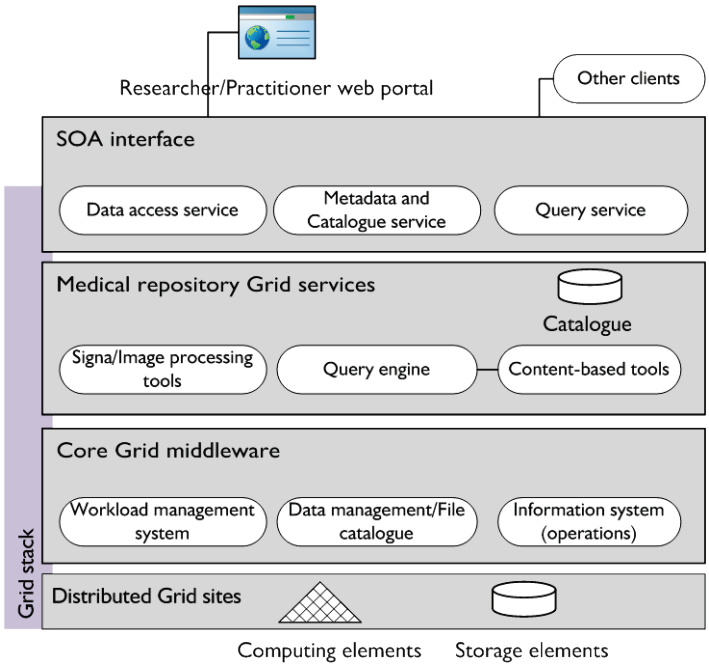


Figure 7.7: Logical architecture for a Grid-enabled medical repository system.

etc.). This portal has been prototyped in our previous work (Oliveira *et al.*, 2009) for the neuroscience research community (Figure 7.9).

2. **SOA Interface**. The service interface is a collection of SOA services providing a programming interface to the repository management using Web Services. These services, expressed at the application-level, include new data upload and registry, data management, standard and customized data query, standard and customized data processing, repository maintenance, etc. The implementation of these services would delegate the heavy work on corresponding Grid functions enabled by a robust middleware.

3. **Medical repository Grid services**. This is the functional core of the system. It will ensure the data workflow management, namely the data processing workflows, supporting the high-level services and the access to basic Grid infrastructure. The Signal/Image Processing tools, Content-based tools and Query Engine analyze data, metadata, system status and deal with operation requests in order to retrieve, organize, store, fetch, and/or process data in a successful way. The **Signal/Image Processing Module** manages the image processing workflows and methods to be used in feature extraction of signal and image information. It is responsible for the identification of suitable Grid methods, monitor their execution, collecting the results and making them available for the requesting entity. The **Content-based tools** are **r**esponsible for feature extraction, either for storing it on the system catalogue or for content based query support. It will be responsible for integrating and store textual and feature information in the catalogue (includes coordinating image processing stages for feature extraction and related textual/clinical inputs). The **Query Engine** manages the access to descriptive data in the catalogue and the access to the actual data/imaging repositories. The **Catalogue** maintains a rich and consistent description of all the data for easy querying. This description includes Grid location, textual information and essential features extracted for imaging sequences (when possible).

4. **Grid Infrastructure**. The Grid infrastructure will support the system layers described above providing seamless computing power, data storage and interconnection between the various subsystems and resources. In the prototype experiments conducted in our group (*e.g.* (Andrade *et al.*, 2007; Oliveira *et al.*, 2009)), the medical data management services were built on top of the gLite middleware.

Functional workflows

We identified three main processes in which most of the functional workflow in the system rely: Data input, Simple Query, Query and Retrieve. These processes can be invoked in isolation or, more naturally, combined to implement use cases. For instance, an assessment may imply a query to the catalogue (Simple Query) which provides partial information, and then, in a second step, a request to get complementary details (Query and Retrieve). Data input process will be responsible for storing the data in the Repository, while ensuring that all relevant features are

extracted and catalogued. The feature extraction may include signal or imaging processing (using Grid as computational resource).

In our model, we assume that the result of a simple query is a list of references and not the actual data, given the size of some volumes. To satisfy a query, the system must analyse the parameters, which, in this case, may involve pre-processing for feature extraction, and transform the query to match the semantics of the Catalogue.

The Query-Retrieve process may be seen as a simple query process followed by the actual access and retrieval of the data volumes.

Study results

The proposed vision for the set-up of a science-friendly repository of medical cases has been prototyped in smaller sub-projects. An extensive amount of effort not reported in this work was oriented to the setup of the enabling Grid infrastructure (Cunha *et al.*, 2007). The local Grid site, integrated in the IberGrid, is a modest contributor to the European Grid Infrastructure, already active since 2007 (Figure 7.8).

The portal system has been successfully deployed in the context of the Brain Imaging Network Grid initiative (BIN-G), led by our research group (Oliveira *et al.*, 2009), providing a user-friendly environment for the remote submission of demanding analysis tasks (Figure 7.9). The portal was tested with data and use cases from the brain imaging and endoscopic capsule communities. Each domain originated specific Grid methods. In the case of the brain imaging, we used the Grid to provide the computational power to calculate association measures between the brain activity at a specific voxel and all the others, using a non-parametric approach (Andrade *et al.*, 2007). The results are presented as association maps, showing the positive or negative association (Figure 7.10).

In the case of the endoscopic video, as discussed in section 7.3.1, the desktop application was prepared to interface with the Grid middleware and allow the remote execution of the costly part of the Automatic Topographic Segmentation algorithm (Oliveira *et al.*, 2010). The user accesses the Grid facilities with a single click in the corresponding menu option. The results are inserted in the CapView visualization, tagging the anatomic regions of interest (Figure 7.5). In a complementary direction, we have also developed a forum to discuss clinical cases online, in a controlled community of experts and learners, using the endoscopic videos for the gastroenterology community (http://www.capview.org/).

These partial developments confirmed that the research infrastructures enabled by the Grid can be used to store medical images and run complex analysis, but additional developments are necessary to integrate the partial components into a coherent solution.

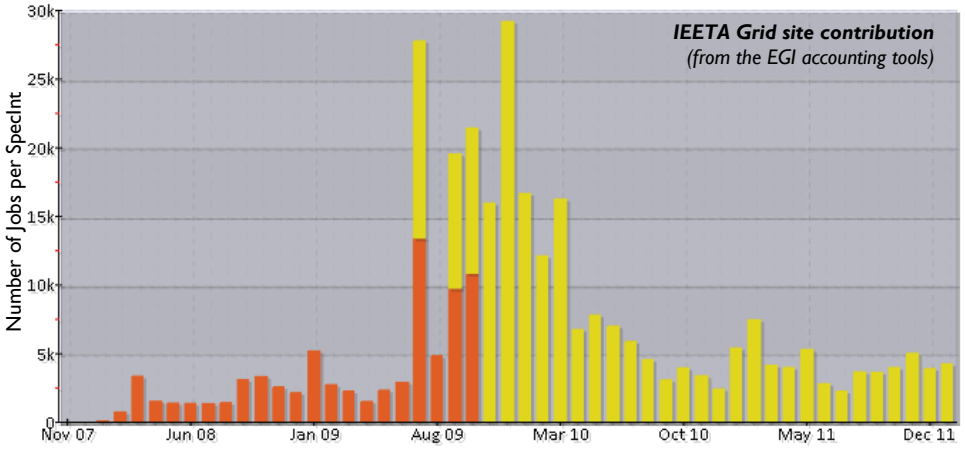


Figure 7.8: SpectInt history of the IEETA Grid site (available from http://accounting.egi.eu/; the change of colours corresponds to a migration of the infrastructure).

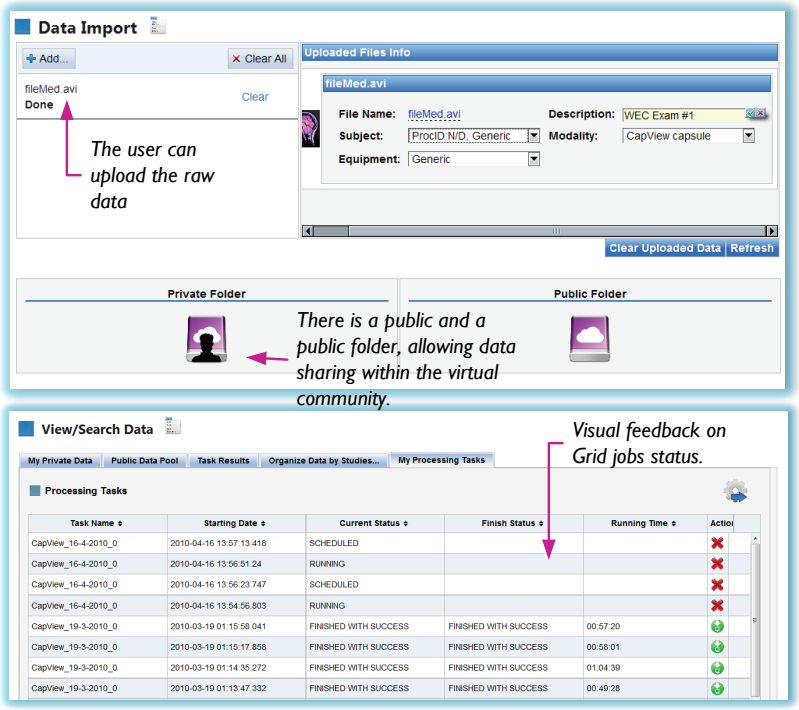


Figure 7.9: The Medical Applications Grid Interface portal for biomedical researchers. The portal allows data uploading and sharing, and the remote submission of computational tasks.

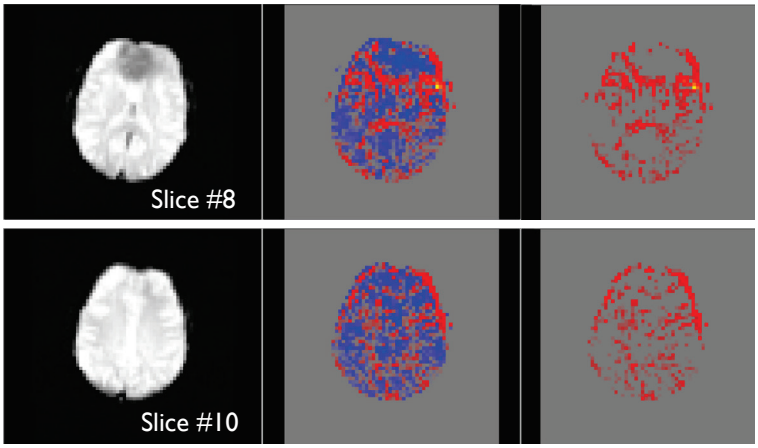


Figure 7.10: fMRI analysis results, represented as association maps between a localized voxel (not shown) and two possible slices. Warm colours represent positive association and blue tones negative.

8 Conclusions

The digital flow of health information is generally seen as insufficient to match the needs of an efficient and safe care delivery. Portugal clearly lacks on routine health information exchange. This can be linked to the absence of a clear strategy and a regulation framework for health information technology, promoting sustained clinical data sharing practices.

In this context, regional initiatives, such as the one presented in this work, can act as incubators of possible solutions for information sharing infrastructures. There are sound arguments to work at the region level, starting with the fact that health care *is* regional, in the sense that there is a strong locality of the patient encounters and cross-institution collaborations, corresponding to the referral pathways defined. In addition, the regional scale provides a more controlled environment in which the introduction of socio-technical changes are best managed.

In this work, we argue that it is possible to enhance the information availability at the point of care in a region without disrupting current practices, systems and responsibilities. This is achieved by a new software layer, the RTSys, working on the integration of heterogeneous decentralized information sources, scattered in the region, to build the abstraction of a virtual EHR. RTSys has been successfully deployed in the region of Aveiro to build the RTS network, connecting eight different care institutions, sharing the virtual EHR of +350.000 inhabitants. The access to the virtual, regional EHR is available at any service point in the partner institutions.

RTSys was designed to adapt to new configurations, especially to allow a dynamic composition of the information sources collection. It is possible for systems to enter in any stage and their information will be included in the discovery processes from that point onward.

The RTS network, supported by RTSys, was able to surpass some traditional and long lasting barriers. It is now possible to have the GPs directly accessing information produced in the Hospitals concerning their patients, or having staff from a Hospital accessing results of diagnostic procedures performed in a different one. The RTS was a novel work in the Portuguese reality, tearing walls down and having organizations opening their data to the others in cooperative workflows. Paraphrasing N. Armstrong, it is a small step for *technology*, a giant leap for *collaborative care*, given the existing context. The opportunity for this kind of approaches was confirmed by an innovation prize awarded to the RTS project.

The active participation of health practitioners in the specification of the ICT functions was, for the most of them, a first-time experience of fruitful collaboration with the engineering teams

in process redesign. This opportunity for mutual growth is often neglected in the Portuguese health system.

The introduction of health information technology in health settings is not easy and bears risks. In this work we took the risk of facing the health care system idiosyncrasies and propose a new collaboration tool, knowing that, ultimately, sharing technology could be overruled by distrust or 'protective' motivations. We had the luck to work with enthusiastic health care professionals that confirmed our roadmap, but the top-level support to the full realization of the RTSys potential, in the context of the RTS network, was not so keen.

It is rewarding to have developed a solution that is not only an academic asset, but one that made its way into the field. In addition, the positive feedback from health professionals allows us to believe that we did something with value for patient data integration in support of their practice.

In a complementary line of work, we have explored the Grid computing technology to enable the wide-scale and integrative management of medical data. From the grand vision of a 'health Grid' connected with RTSys to enable research workflows, we have proactively developed and deployed partial components. The particular scenario of brain imaging tools for neuroscience has been explored to demonstrate the feasibility of the Grid approach, originating the development of a domain-specific portal for the brain research community. Using the portal, end-users can upload data (medical images) to the Grid and run specific operators. The use of the Grid as a medical imaging storage and processing tool integrated with a desktop application was demonstrated for the analysis of endoscopic videos.

Opportunities for future work

Having gone through all the phases of ICT solutions development, from requirements engineering to system introduction in the health care practice, there are some opportunities for future work to enhance the RTSys solution. An important line of work would be the enhancement of semantic models. The present content model is a bottom-up effort, based on the analysis of the existing reality; it is crystallized in the software entities comprising the solution and it is not possible for domain experts to expand the existing representations. The adoption of a two-level model could leverage the ability to manage domain knowledge and software processes independently. This support, extended with the introduction of semantic annotation of sources, could also enable automatic information integration through semantic mediation for other applications besides the R-EHR.

The RTSys adopts a read-only information bus approach; it does not support the definition of shared care plans, for example. This presents an opportunity for a possible evolution of the system, towards easier access to the clinical context of the patient, in line with the central objective of our work. In the care plans use case, it would be possible for a professional to define a therapeutic plan for a Patient and have it shared (and contributed) by all relevant actors in the network. The shared care plan can be linked to the distributed clinical information systems, relating the plan with the information in the operational systems.

RTSys implements features that would be relevant to support a Personal Health Record connected to the operational clinical information systems. It would be easy to add new patient-oriented services to the platform, allowing the access to more information, a richer interaction with the health actors in the region, and the recording of patient contributed clinical content (though this discussion is much beyond a technical issue).

Another line of work is naturally the development of the ideas outlined in Chapter 7, towards the full realization of a HealthGrid connected with RTSys, to bridge the care information infrastructure and e-Science infrastructures. After having developed a theoretical approach and experimental prototypes, new joint work with the domain experts is required to define the most promising use cases on the RTSys/Grid approach and establish a roadmap towards their implementation. We believe that future health information technology will explore this paradigm, in which clinical tools and biomedical resources will be used together, benefiting from global, scalable and secure computing infrastructures.

9 References

Adler-Milstein, J., Bates, D. W. & Jha, A. K. (2011a). A survey of health information exchange organizations in the united states: Implications for meaningful use. *Annals of Internal Medicine*, 154(10), pp. 666-671.

Adler-Milstein, J., DesRoches, C. M. & Jha, A. K. (2011b). Health information exchange among US hospitals. *American Journal of Managed Care*, 17(11), pp. 761-768.

Adler-Milstein, J., McAfee, A. P., Bates, D. W. & Jha, A. K. (2008). The state of regional health information organizations: Current activities and financing. *Health Affairs*, 27(1), pp. 60-69.

Alonso, G., Casati, F., Kuno, H. & Machiraju, V. (2004). *Web Services: Concepts, Architectures and Applications*, Springer.

Alpdemir, M., Mukherjee, A., Gounaris, A., Paton, N., Fernandes, A., Sakellariou, R., Watson, P. & Li, P. (2005). Using OGSA-DQP to Support Scientific Applications for the Grid. *In:* Herrero, P., Pérez, M. & Robles, V. (eds.) *Scientific Applications of Grid Computing*. Springer Berlin / Heidelberg.

Altman, R. B. (1998). Bioinformatics in support of molecular medicine. *Proceedings / AMIA ... Annual Symposium*, pp. 53-61.

Alur, D., Crupi, J. & Malks, D. (2003). *Core J2EE Patterns: Best Practices and Design Strategies*, Sun Microsystems Press.

Anderson, J. G. (2007). Social, ethical and legal barriers to E-health. *International Journal of Medical Informatics*, 76(5–6), pp. 480-483.

Andoulsi, I., Blanquer, I., Breton, V., Dobrev, A., Doosselaere, C. V., Hernandez, V., Herveg, J., Jacq, N., Legré, Y., Olive, M., *et al.* (2008). *SHARE Integrated Road Map II*, SHARE Project (SHARE Deliverable D6.2). Available: http://roadmap.healthgrid.org/.

Andrade, R., Oliveira, I., Fernandes, J. M. & Cunha, J. P. (2007). A Grid framework for non-linear brain FMRI analysis. *Studies in health technology and informatics*, 126, pp. 299-305.

Anonymous (1998). Law no. 67/98, from October 26: Law on Personal data protection. The Portuguese Assembly of the Republic.

Armbrust, M., Fox, A., Griffith, R., Joseph, A. D., Katz, R., Konwinski, A., Lee, G., Patterson, D., Rabkin, A., Stoica, I., *et al.* (2010). A view of cloud computing. *Commun. ACM*, 53(4), pp. 50-58.

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., Davis, A. P., Dolinski, K., Dwight, S. S., Eppig, J. T., *et al.* (2000). Gene Ontology: tool for the unification of biology. *Nature Genetics*, 25(25-29).

Aspden, P., Corrigan, J. M., Wolcott, J. & Erickson, S. M. (eds.) (2004). *Patient Safety : Achieving a New Standard for Care,* Washiton. D.C.: The National Academies Press.

ASTM-E31.25 (2012). Standard Specification for Continuity of Care Record (CCR). ASTM International. ASTM E2369.

Barros, P. P. & Simões, J. d. A. (2007). Portugal: Health system review. *In:* Allin, S. & Mossialos, E. (eds.) *Health Systems in Transition.* European Observatory on Health Systems and Policies. 9.

Bates, D. W. & Gawande, A. A. (2003). Improving Safety with Information Technology. *New England Journal of Medicine*, 348(25), pp. 2526-2534.

Baumlin, K. M., Genes, N., Landman, A., Shapiro, J. S., Taylor, T. & Janiak, B. (2010). Electronic collaboration: Using technology to solve old problems of quality care. *Academic Emergency Medicine*, 17(12), pp. 1312-1321.

Beale, T. (2002). Archetypes: Constraint-based domain models for future-proof information systems. *Workshop on Behavioural Semantics (OOPSLA'02)*. Seatle, WA.

Bellotti, R., Cerello, P., Tangaro, S., Bevilacqua, V., Castellano, M., Mastronardi, G., De Carlo, F., Bagnasco, S., Bottigli, U., Cataldo, R., *et al.* (2007). Distributed medical images analysis on a Grid infrastructure. *Future Generation Computer Systems*, 23(3), pp. 475-484.

Benkner, S., Arbona, A., Berti, G., Chiarini, A., Dunlop, R., Engelbrecht, G., Frangi, A. F., Friedrich, C. M., Hanser, S., Hasselmeyer, P., *et al.* (2010). @neurIST: Infrastructure for advanced disease management through integration of heterogeneous data, computing, and complex processing services. *IEEE Transactions on Information Technology in Biomedicine*, 14(6), pp. 1365-1377.

Berg, M. (2001). Implementing information systems in health care organizations: Myths and challenges. *International Journal of Medical Informatics*, 64(2-3), pp. 143-156.

Berg, M. (ed.) (2004). *Health information management: integrating information technology in health care work*, New York: Routledge.

Bergmann, J., Bott, O. J., Pretschner, D. P. & Haux, R. (2007). An e-consent-based shared EHR system architecture for integrated healthcare networks. *International Journal of Medical Informatics*, 76(2-3), pp. 130-136.

Berners-Lee, T., Hendler, J. & Lassila, O. (2001). The semantic web. *Scientific American*, 284(5), pp. 34-43.

Bird, I., Jones, B. & Kee, K. F. (2009). The Organization and Management of Grid Infrastructures. *Computer*, 42(1), pp. 36-46.

Bisbal, J. & Berry, D. (2011). An analysis framework for electronic health record systems: Interoperation and collaboration in shared healthcare. *Methods of Information in Medicine*, 50(2), pp. 180-189.

Blanquer Espert, I., Hernández García, V., Meseguer Anastásio, F. J. & Segrelles Quilis, J. D. (2009). Content-based organisation of virtual repositories of DICOM objects. *Future Generation Computer Systems*, 25(6), pp. 627-637.

Blanquer, I., Hernández, V., Salavert, J. & Segrelles, D. (2010). Integrating TRENCADIS components in gLite to share DICOM medical images and structured reports. *8th HealthGrid Conference, HealthGrid 2010*. Orsay: IOS Press. 159: 64-75.

Blind, K. (2004). *The Economics Of Standards: Theory, Evidence, Policy*, Edward Elgar.

Blobel, B., Engel, K. & Pharow, P. (2006). Semantic interoperability -- HL7 Version 3 compared to advanced architecture standards. *Methods of Information in Medicine*, 45(4), pp. 343-353.

Blobel, B. & Holena, M. (1997). Comparing middleware concepts for advanced healthcare system architectures. *International Journal of Medical Informatics*, 46(2), pp. 69-85.

Bloch, J. (2008). *Effective Java*, Addison-Wesley.

Blumenthal, D. (2009). Stimulating the adoption of health information technology. *New England Journal of Medicine*, 360(15), pp. 1477-1479.

Blumenthal, D. & Tavenner, M. (2010). The "meaningful use" regulation for electronic health records. *New England Journal of Medicine*, 363(6), pp. 501-504.

Bodenreider, O. & Stevens, R. (2006). Bio-ontologies: Current trends and future directions. *Briefings in Bioinformatics*, 7(3), pp. 256-274.

Booch, G., Rumbaugh, J. & Jacobson, I. (2005). *The Unified Modelling Language User Guide*, Addison-Wesley.

Breton, V., Dean, K., Solomonides, T., Blanquer, I., Hernandez, V., Medico, E., Maglaveras, N., Benkner, S., Lonsdale, G., Lloyd, S., *et al.* (2005). The Healthgrid White Paper. *Studies in health technology and informatics*, 112, pp. 249-321.

Bui, A. A. T., Aberle, D. R. & Kangarloo, H. (2007). TimeLine: Visualizing integrated patient records. *IEEE Transactions on Information Technology in Biomedicine*, 11(4), pp. 462-473.

Burges, C. J. C. (1998). A tutorial on support vector machines for pattern recognition. *Data Mining and Knowledge Discovery*, 2(2), pp. 121-167.

Cannataro, M. & Talia, D. (2004). Semantics and Knowledge Grids: Building the Next-Generation Grid. *IEEE Intelligent Systems*, 19(1), pp. 56-63.

Cannataro, M., Veltri, P., Aloisio, G., Fiore, S., Mirto, M., Maglaveras, N., Chouvarda, I., V. Koutkias, Malousi, A., Lopez, V., *et al.* (2005). Genomic Medicine Grid. *In:* Breton, V., Dean, K. & Solomonides, T. (eds.) *The Healthgrid White Paper.*

Chang, S. F., Sikora, T. & Puri, A. (2001). Overview of the MPEG-7 standard. *IEEE Transactions on Circuits and Systems for Video Technology*, 11(6), pp. 688-695.

Chaudhry, B., Wang, J., Wu, S. Y., Maglione, M., Mojica, W., Roth, E., Morton, S. C. & Shekelle, P. G. (2006). Systematic review: Impact of health information technology on quality, efficiency, and costs of medical care. *Annals of Internal Medicine*, 144(10), pp. 742-752.

Chen, R. (2009). *Towards Interoperable and Knowledge-Based Electronic Health Records Using Archetype Methodology.* PhD thesis, Linköping University.

Chen, R., Klein, G., Sundvall, E., Karlsson, D. & Ahlfeldt, H. (2009). Archetype-based conversion of EHR content models: pilot experience with a regional EHR system. *BMC Medical Informatics and Decision Making*, 9(1), p. 33.

CNECV (2011). *Parecer n.º 60 do Conselho Nacional da Ética para as Ciências da Vida. Parecer sobre informação de saúde e registos informáticos de saúde [The National Council for Ethics in Life Sciences: ruling on health information and electronic health records]*, Conselho Nacional de Ética para as Ciências da Vida (CNECV) (Parecer 60/CNECV/2011).

CNPD (2004). *Relatório de auditoria ao tratamento de informação de saúde nos hospitais [Auditing report on the health information processing in Hospitals]*, Comissão Nacional de Protecção de Dados (CNPD).

CoE (1981). Convention 108: Convention for the Protection of Individuals with regard to Automatic Processing of Personal Data. Council of Europe.

Collins, F. S. & McKusick, V. A. (2001). Implications of the Human Genome Project for medical science. *JAMA*, (285), pp. 540-4.

CRISP (2009). *The Chesapeake Regional Information System for our Patients (CRISP) : A Plan for Statewide Health Information Exchange in Maryland*, Maryland Health Care Commission. Available: http://mhcc.dhmh.maryland.gov/hit/hiePolicyBoard/Documents/CRISP_FinalReport.pdf.

Cronin, C. (2004). *Patient-Centered Care : An Overview of Definitions and Concepts*, Prepared for the National Health Council.

Cruz-Correia, R. (2008). *From Patient Data to Patient Care: Implementation, monitoring and utilization of an integrated Hospital Information System.* PhD Thesis, Universidade do Porto.

Cruz-Correia, R., Vieira-Marques, P., Costa, P., Ferreira, A., Oliveira-Palhares, E., Araújo, F. & Costa-Pereira, A. (2005). Integration of hospital data using agent technologies – A case study. *AI Communications*, 18(3), pp. 191-200.

Cruz-Correia, R., Vieira-Marques, P., Ferreira, A., Almeida, F., Wyatt, J. & Costa-Pereira, A. (2007). Reviewing the integration of patient data: how systems are evolving in practice to meet patient needs. *BMC Medical Informatics and Decision Making*, 7(1), p. 14.

Cruz, I., Cunha, J. P., Mano, L., Polónia, D., Oliveira, I. C. & Soares, P. (2005). *Relatório de Análise de Processos e Fluxos de Informação [Report on processes and infomation flow analysis]*, Projeto Rede Telemática da Saúde.

Cruz, I. A., Silva, A. G., Soares, P., Oliveira, I., Serrano, J. A. V. & Cunha, J. P. S. (2002). Modelling the hospital into the future with UML. *International Journal of Healthcare Technology and Management*, 4(2-4), pp. 193-204.

Cunha, J. P. S., Coimbra, M., Campos, P. & Soares, J. M. (2008). Automated topographic segmentation and transit time estimation in endoscopic capsule exams. *IEEE Transactions on Medical Imaging*, 27(1), pp. 19-27.

Cunha, J. P. S., Oliveira, I., Fernandes, J. M., Campilho, A., Castelo-Branco, M., Sousa, N. & Pereira, A. S. (2007). *BING: The Portuguese Brain Imaging Network Grid,* Santiago Compostella, Cesga-Centro Supercomputacion Galicia.

Cunha, M. B., Cunha, J. P. S. & Silva, T. O. (1997). SIGIF: A digital signal interchange format with application in neurophysiology. *IEEE Transactions on Biomedical Engineering*, 44(5), pp. 413-418.

De Roure, D., Jennings, N. R. & Shadbolt, N. R. (2005). The Semantic Grid: Past, present, and future. *Proceedings of the IEEE*, 93(3), pp. 669-680.

Deelman, E., Gannon, D., Shields, M. & Taylor, I. (2009). Workflows and e-Science: An overview of workflow system features and capabilities. *Future Generation Computer Systems*, 25(5), pp. 528-540.

Dennis, G. (2005). Canada Health Infoway - Towards a National Interoperable Electronic Health Record (EHR) Solution. *Regional Health Economies and ICT Services: The PICNIC Experience.* IOS Press.

Detmer, D., Bloomrosen, M., Raymond, B. & Tang, P. (2008). Integrated personal health records: Transformative tools for consumer-centric care. *BMC Medical Informatics and Decision Making*, 8.

Deutsch, E., Duftschmid, G. & Dorda, W. (2010). Critical areas of national electronic health record programs—Is our focus correct? *International Journal of Medical Informatics*, 79(3), pp. 211-222.

DH+RCGP+BMA (2011). *The Good Practice Guidelines for GP electronic patient records - version 4 (2011)*, Department of Health (DH)/Royal College of General Practitioners (RCGP)/British Medical Association (BMA). Available: http://www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_125310.

Dias, G. S., Oliveira, J. L., Vicente, J. & Martin-Sanchez, F. (2006). Integrating medical and genomic data: a successful example for rare diseases. *Studies in health technology and informatics*, 124, pp. 125-130.

Dick, R. S., Steen, E. B. & Detmer, D. E. (eds.) (1997). *The Computer-Based Patient Record: An Essential Technology for Health Care, Revised Edition,* Washington, D.C.: The National Academies Press.

Dix, A., Finlay, J. & Abowd, G. D. (2004). *Human-computer interaction*, Addison-Wesley.

Dobrev, A., Haesner, M., Hüsing, T., Korte, W. B. & Meyer, I. (2008). *Benchmarking ICT use among General Practitioners in Europe*, European Comission. Available: http://ec.europa.eu/information_society/eeurope/i2010/docs/benchmarking/gp_survey_final_report.pdf.

Dobrev, A., Jones, T., Stroetmann, K., Vatter, Y. & Peng, K. (2009). *The socio-economic impact of interoperable electronic health record (EHR) and ePrescribing systems in Europe and beyond (Final study report)*, EHR-Impact Project.

Dobrev, A., Jones, T., Stroetmann, V., Stroetmann, K., Vatter, Y. & Peng, K. (2010). *Interoperable eHealth is Worth it: Securing Benefits from Electronic Health Records and ePrescribing*, European Comission. Available: http://www.ehr-impact.eu/downloads/documents/EHRI_final_report_2009.pdf.

Dolin, R. H., Alschuler, L., Boyer, S., Beebe, C., Behlen, F. M., Biron, P. V. & Shabo, A. (2006). HL7 clinical document architecture, release 2. *Journal of the American Medical Informatics Association*, 13(1), pp. 30-39.

Donnelly, J., Mussi, J., Parisot, C. & Russler, D. (2006). Building an interoperable regional health information network today with IHE integration profiles. *Journal of healthcare information management : JHIM*, 20(3), pp. 29-38.

Doosselaere, C. v., Herveg, J., Silber, D. & Wilson, P. (2008). *Legally eHealth - Putting eHealth in its European legal context*, European Comission - Information Society and Media. Available: http://ec.europa.eu/information_society/activities/health/docs/studies/legally_ehealth/legally-ehealth-report.pdf.

EC-SHARE (2008). *SHARE the journey - A European Healthgrid Roadmap*, European Comission - Information Society and Media. Available: http://ec.europa.eu/information_society/activities/health/docs/publications/200810share-roadmap.pdf.

EC (2004). Communication from the Commission to the Council, the European Parliament, the European Economic and Social Committee and the Committee of the Regions - e-Health - making healthcare better for European citizens: an action plan for a European e-Health Area. *COM (2004) 356.*

EC (2006). *Connected Health: Quality and Safety for European Citizens*, European Comission - Unit ICT for Health. Available: http://ec.europa.eu/information_society/activities/health/docs/policy/connected-health_final-covers18092006.pdf.

EC (2007a). *Accelerating the Development of the eHealth Market in Europe*, European Comission - eHealth Taskforce report 2007. Available: http://ec.europa.eu/information_society/activities/health/docs/publications/lmi-report-final-2007dec.pdf.

EC (2007b). *eHealth for Safety - Impact of ICT on Patient Safety and Risk Management* Available: http://ec.europa.eu/information_society/activities/health/downloads/index_en.htm.

EC (2007c). *eHealth priorities and strategies in European countries.* Available: http://ec.europa.eu/information_society/activities/health/docs/policy/ehealth-era-full-report.pdf.

EC (2008a). Commission Recommendation of 2 July 2008 on cross-border interoperability of electronic health record systems. *C(2008)3282*

EC (2008b). Communication from the Commission to the European Parliament, the Council, the European Economic and Social Committee and the Committee of the Regions on telemedicine for the benefit of patients, healthcare systems and society. *COM(2008)689 final.*

EC (2009). *eHealth in Europe.* Available: http://www.ehealtheurope.net/img/document_library0282/eHealthasolutionforEuropeanhealthcaresystems%5B1%5D.pdf.

Edwards, A., Hollin, I., Barry, J. & Kachnowski, S. (2010). Barriers to cross--institutional health information exchange: a literature review. *Journal of healthcare information management : JHIM*, 24(3), pp. 22-34.

Eichelberg, M., Aden, T., Riesmeier, J., Dogac, A. & Laleci, G. B. (2005). A survey and analysis of Electronic Healthcare Record standards. *ACM Computing Surveys*, 37(4), pp. 277-315.

Elger, B. S., Iavindrasana, J., Lo Iacono, L., Müller, H., Roduit, N., Summers, P. & Wright, J. (2010). Strategies for health data exchange for secondary, cross-institutional clinical research. *Computer Methods and Programs in Biomedicine*, 99(3), pp. 230-251.

Empirica (2008). *ICT standards in the health sector: current situation and prospects*, European Comission - DG Enterprise & Industry. Available: http://ec.europa.eu/enterprise/archives/e-business-watch/studies/special_topics/2007/documents/Special-study_01-2008_ICT_health_standards.pdf.

epSOS (2010). D3.2.2 Final definition of functional service requirements - Patient Summary. Available: http://www.epsos.eu/home/download-area/deliverables.html [Accessed 2012-02-23].

epSOS (2011). *epSOS Project* [Online]. Available: http://www.epsos.eu [Accessed 2012-02-23].

Erberich, S. G., Silverstein, J. C., Chervenak, A., Schuler, R., Nelson, M. D. & Kesselman, C. (2007). Globus MEDICUS - federation of DICOM medical imaging devices into healthcare Grids. *Studies in health technology and informatics*, 126, pp. 269-278.

Erl, T. (2004). *Service-Oriented Architecture: A Field Guide to Integrating XML and Web Services*, Prentice Hall PTR.

Espanha, R. (2010). Tecnologias de Informação e Comunicação na Saúde. *In:* Simões, J. (ed.) *30 Anos do Serviço Nacional de Saúde - Um Percurso Comentado.* Almedina.

European Union (1995). Directive 95/46/EC of the European Parliament and of the Council of 24 October 1995 on the protection of individuals with regard to the processing of personal data and on the free movement of such data. *Directive 95/46/EC*.

European Union (2011). Directive 2011/24/EU of the European Parliament and of the Council of 9 March 2011 on the application of patients' rights in cross-border healthcare. *Directive 2011/24/EU*.

Featherstone, I. & Keen, J. (2012). Do integrated record systems lead to integrated services? An observational study of a multi-professional system in a diabetes service. *International Journal of Medical Informatics*, 81(1), pp. 45-52.

Ferranti, J. M., Musser, R. C., Kawamoto, K. & Hammond, W. E. (2006). The Clinical Document Architecture and the Continuity of Care Record: A Critical Analysis. *Journal of the American Medical Informatics Association*, 13(3), pp. 245-252.

Ferrara, F. (1998). The standard 'Healthcare Information Systems Architecture' and the DHE middleware. *International Journal of Medical Informatics*, 52(1-3), pp. 39-51.

Ferreira, A., Cruz-Correia, R., Antunes, L., Farinha, P., Oliveira-Palhares, E., Chadwick, D. W. & Costa-Pereira, A. (2006). How to Break Access Control in a Controlled Manner. *19th IEEE International Symposium on Computer-Based Medical Systems 2006 (CBMS 2006)*. Salt Lake City, Utah.

Ferreira, L. (2011). *Medical Information Extraction in European Portuguese*. PhD, University of Aveiro.

Ferreira, L., Teixeira, A. & Cunha, J. P. S. (2012). *Medical Information Extraction: Information Extraction from Portuguese Hospital Discharge Letters*, LAP LAMBERT Academic Publishing.

Florescu, D., Levy, A. & Mendelzon, A. (1998). Database techniques for the World-Wide Web: a survey. *SIGMOD Rec.*, 27(3), pp. 59-74.

Fontaine, P., Ross, S. E., Zink, T. & Schilling, L. M. (2010). Systematic review of health information exchange in primary care practices. *Journal of the American Board of Family Medicine*, 23(5), pp. 655-670.

Foster, I., Kesselman, C., Nick, J. M. & Tuecke, S. (2003). The Physiology of the Grid. *Grid Computing*. John Wiley & Sons, Ltd.

Foster, I., Kesselman, C. & Tuecke, S. (2001). The Anatomy of the Grid: Enabling scalable virtual organizations. *Internation Journal of High Performance Computing Applications*, 15(3), pp. 200-222.

Foster, I., Yong, Z., Raicu, I. & Lu, S. (2008). Cloud Computing and Grid Computing 360-Degree Compared. Grid Computing Environments Workshop 2008 (GCE '08), 12-16 Nov 2008. pp. 1-10.

Freund, J., Comaniciu, D., Ioannis, Y., Liu, P., McClatchey, R., Morley-Fletcher, E., Pennec, X., Pongiglione, G. & Zhou, X. S. (2006). Health-e-child: an integrated biomedical platform for grid-based paediatric applications. *Studies in health technology and informatics*, 120, pp. 259-270.

Frisse, M. E. & Holmes, R. L. (2007). Estimated financial savings associated with health information exchange and ambulatory care referral. *Journal of Biomedical Informatics*, 40(6 SUPPL.), pp. S27-S32.

Gamma, E., Helm, R., Johnson, R. & Vlissides, J. (1995). *Design Patterns: Elements of Reusable Object-Oriented Software*, Addison-Wesley.

Garde, S., Knaup, P., Hovenga, E. J. S. & Heard, S. (2007). Towards semantic interoperability for electronic health records - Domain knowledge governance for openEHR archetypes. *Methods of Information in Medicine*, 46(3), pp. 332-343.

Geddes, N. (2012). The Large Hadron Collider and Grid computing. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 370(1961), pp. 965-977.

Gomes, H., Cunha, J. P. S. & Zuquete, A. (2007). Authentication architecture for eHealth professionals. *In:* Meersman, R. & Tari, Z. (eds.) *On the Move to Meaningful Internet Systems 2007: CoopIS, DOA, ODBASE, GADA, and IS, Proceedings*. Berlin: Springer-Verlag Berlin.

Goroll, A. H., Simon, S. R., Tripathi, M., Ascenzo, C. & Bates, D. W. (2009). Community-wide Implementation of Health Information Technology: The Massachusetts eHealth Collaborative Experience. *J Am Med Inform Assoc*, 16(1), pp. 132-139.

Gray, J., Liu, D. T., Nieto-Santisteban, M., Szalay, A., Dewitt, D. J. & Heber, G. (2005). Scientific data management in the coming decade. *SIGMOD Record*, 34(4), pp. 34-41.

Greenhalgh, T., Stramer, K., Bratan, T., Byrne, E., Russell, J., Hinder, S. & Potts, H. (2010a). *THE DEVIL'S IN THE DETAIL : Final report of the independent evaluation of the Summary Care Record and HealthSpace programmes*, University College London.

Greenhalgh, T., Stramer, K., Bratan, T., Byrne, E., Russell, J. & Potts, H. W. W. (2010b). Adoption and non-adoption of a shared electronic summary record in England: A mixed-method case study. *BMJ*, 340(7761), p. 1399.

Grimson, J., Grimson, W. & Hasselbring, W. (2000). The SI challenge in health care. *Commun. ACM*, 43(6), pp. 48-55.

Grimson, W., Jung, B., Van Mulligen, E. M., Van Ginneken, A., Pardon, S. & Sottile, P. A. (2002). Extensions to the HISA standard - The SynEX computing environment. *Methods of Information in Medicine*, 41(5), pp. 401-410.

Gupta, A., Ludäscher, B. & Martone, M. (2003). Registering Scientific Information Sources for Semantic Mediation. *In:* Spaccapietra, S., March, S. & Kambayashi, Y. (eds.) *Conceptual Modeling - ER 2002 Proceedings*. Springer Berlin / Heidelberg.

Gwadry-Sridhar, F., Olabarriaga, S. D., Van Kampen, A., Bauer, M. & Solomonides, T. (2010). Hitting the ground running: Healthgrid deployment and adoption. 8th HealthGrid Conference (HealthGrid 2010), 28-30 June 2010 Orsay. Studies in Health Technology and Informatics, 40-51.

Halamka, J., Aranow, M., Ascenzo, C., Bates, D., Debor, G., Glaser, J., Goroll, A., Stowe, J., Tripathi, M. & Vineyard, G. (2005). Health care IT collaboration in Massachusetts: The experience of creating regional connectivity. *Journal of the American Medical Informatics Association*, 12(6), pp. 596-601.

Halevy, A. Y. (2001). Answering queries using views: A survey. *Vldb Journal*, 10(4), pp. 270-294.

Hammond, W. E. & Cimino, J. J. (2006). Standards in Biomedical Informatics. *In:* Shortliffe, E. H. & Cimino, J. J. (eds.) *Biomedical Informatics*. Springer New York.

Harno, K. & Ruotsalainen, P. (2006). Sharable EHR systems in Finland. *Studies in health technology and informatics*, 121, pp. 364-370.

Hastings, S., Gray, M., Catalyurek, U., Gingrich, T., Young, D. & Saltz, J. (2002). Grid Support for Collaborative Clinical and Biomedical Research Studies. Imaging and the Internet 2002, 2002 Pittsburgh, Pennsylvania.

Haux, R. (2006). Health information systems - past, present, future. *International Journal of Medical Informatics*, 75(3-4), pp. 268-281.

Häyrinen, K., Saranto, K. & Nykänen, P. (2008). Definition, structure, content, use and impacts of electronic health records: A review of the research literature. *International Journal of Medical Informatics*, 77(5), pp. 291-304.

Helmer, K. G., Ambite, J. L., Ames, J., Ananthakrishnan, R., Burns, G., Chervenak, A. L., Foster, I., Liming, L., Keator, D., Macciardi, F., *et al.* (2011). Enabling collaborative research using the Biomedical Informatics Research Network (BIRN). *Journal of the American Medical Informatics Association*, 18(4), pp. 416-422.

Hey, T., Tansley, S. & Tolle, K. (eds.) (2009). *The Fourth Paradigm: Data-Intensive Scientific Discovery*, Redmond, WA: Microsoft Research.

Hey, T. & Trefethen, A. (2003). e-Science and its implications. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 361(1809), pp. 1809-1825.

HGP *Human Genome Project Information* [Online]. Available: http://www.ornl.gov/sci/techresources/Human_Genome/home.shtml [Accessed 2012-02-23].

HITSP (2008). *C 32 - HITSP Summary Documents Using HL7 Continuity of Care Document (CCD) Component*, Healthcare Information Technology Standards Panel

HL7 (2012a). *Health Level Seven (HL7)* [Online]. Available: http://www.hl7.org/ [Accessed 2010-01-28].

HL7 (2012b). *Health Level Seven Version 3 - Reference Information Model* [Online]. Health Level Seven, Inc. Available: http://www.hl7.org/implement/standards/rim.cfm.

HL7 (2012c). *HL7/ASTM Implementation Guide for CDA Release 2 -Continuity of Care Document (CCD)*, Health Level Seven International

HL7/ISO-TC215 (2009). *Electronic Health Record-System Functional Model, Release 1.1*, ISO/HL7 10781:2009.

HN-SC *Helix Nebula - The Science Cloud* [Online]. Available: http://helix-nebula.eu/ [Accessed 2012-04-01].

Hoerbst, A. & Ammenwerth, E. (2010a). Electronic Health Records A Systematic Review on Quality Requirements. *Methods of Information in Medicine*, 49(4), pp. 320-336.

Hoerbst, A., Kohl, C. D., Knaup, P. & Ammenwerth, E. (2010b). Attitudes and behaviors related to the introduction of electronic health records among Austrian and German citizens. *International Journal of Medical Informatics*, 79(2), pp. 81-89.

Housley, R., Polk, W., Ford, W. & Solo, D. (2002). *Internet X.509 Public Key Infrastructure Certificate and Certificate Revocation List (CRL) Profile.* Available: https://tools.ietf.org/html/rfc3280.

HSSP (2008). *The Practical Guide for SOA in Health Care : A real-world approach to planning, designing, and deploying SOA.*, Healthcare Services Specification Project (a collaborative between Health Level Seven and the Object Management Group). Available: http://hssp.wikispaces.com/PracticalGuide [Accessed 2012-01-20].

HSSP (2012). *Healthcare Services Specification Project* [Online]. Available: http://hssp.wikispaces.com/ [Accessed 2012-02-23].

ICD-CM (2011). *International Classification of Diseases - Clinical Modification 10th Revision* [Online]. Center for Disease Control and Prevention. Available: http://www.cdc.gov/nchs/icd/icd10cm.htm [Accessed 2012-02-23].

ICU (2012). *Informação Clínica do Utente (website)* [Online]. Available: http://epr.med.up.pt/icu/ [Accessed 2012-02-23].

IHE (2012). *IHE Technical Frameworks* [Online]. Available: http://www.ihe.net/Technical_Framework/ [Accessed 2012-01-20].

Infoway (2006). *EHRS Blueprint v2*, Canada Health Infoway. Available: https://www.infoway-inforoute.ca/index.php/resources/technical-documents/doc_download/505-ehrs-blueprint-v2-full.

Institute of Medicine (2001). *Crossing the Quality Chasm : A New Health System for the 21st Century*, Washington, D.C., National Academies Press.

Intel (2012). *Intel SOA Expressway for Healthcare* [Online]. Available: http://www.intel.com/about/companyinfo/healthcare/products/soa/index.htm [Accessed 2012-02-23].

ISO-TC215 (2005). *Health informatics -- Electronic health record -- Definition, scope, and context*, International Organization for Standardization (ISO/TR 20514).

ISO-TC215 (2008). *Health informatics - Electronic health record communication - Part 1: Reference model*, International Organization for Standardization (ISO 13606-1:2008).

ISO-TC215 (2009a). *Health informatics -- Service architecture -- Part 1: Enterprise viewpoint*, International Organization for Standardization (ISO 12967-1:2009).

ISO-TC215 (2009b). *Health informatics -- Service architecture -- Part 2: Information viewpoint*, International Organization for Standardization (ISO 12967-2:2009).

ISO-TC215 (2009c). *Health informatics -- Service architecture -- Part 3: Computational viewpoint*, International Organization for Standardization (ISO 12967-3:2009).

ISO-TC215 (2011). *ISO 18308:2011 Health informatics -- Requirements for an electronic health record architecture*, International Organization for Standardization (ISO 18308:2011).

ISO/IEC (1998). *Information technology -- Open Distributed Processing -- Reference model: Overview*, ISO/IEC 10746-1:1998.

Jensen, H. B. & Pedersen, C. D. (2004). MedCom: Danish Health Care Network. *In:* Iakovidis, I., Wilson, P. & Healy, J. C. (eds.) *E-Health: Current Situation and Examples of Implemented and Beneficial E-Health Applications*. IOS Press.

Jha, A. K., Chan, D. C., Ridgway, A. B., Franz, C. & Bates, D. W. (2009a). Improving Safety And Eliminating Redundant Tests: Cutting Costs In U.S. Hospitals. *Health Aff*, 28(5), pp. 1475-1484.

Jha, A. K., DesRoches, C. M., Campbell, E. G., Donelan, K., Rao, S. R., Ferris, T. G., Shields, A., Rosenbaum, S. & Blumenthal, D. (2009b). Use of Electronic Health Records in U. S. Hospitals. *New England Journal of Medicine*, 360(16), pp. 1628-1638.

Josuttis, N. M. (2007). *SOA in Practice: The Art of Distributed System Design*, O'Reilly.

Kadry, B., Sanderson, I. C. & Macario, A. (2010). Challenges that limit meaningful use of health information technology. *Current Opinion in Anaesthesiology*, 23(2), pp. 184-192.

Kaelber, D. C. & Bates, D. W. (2007). Health information exchange and patient safety. *Journal of Biomedical Informatics*, 40(6 SUPPL.), pp. S40-S45.

Kaelber, D. C., Jha, A. K., Johnston, D., Middleton, B. & Bates, D. W. (2008). A Research Agenda for Personal Health Records (PHRs). *J Am Med Inform Assoc*, 15(6), pp. 729-736.

Kalra, D. (2006). Electronic health record standards. *Yearbook Of Medical Informatics*, pp. 136-144.

Kalra, D., Singleton, P., Milan, J., MacKay, J., Detmer, D., Rector, A. & Ingram, D. (2005). Security and confidentiality approach for the Clinical E-Science Framework (CLEF). *Methods of Information in Medicine*, 44(2), pp. 193-197.

Katehakis, D. G., Sfakianakis, S. G., Kavlentakis, G., Anthoulakis, D. N. & Tsiknakis, M. (2007). Delivering a lifelong integrated electronic health record based on a service oriented architecture. *IEEE Transactions on Information Technology in Biomedicine*, 11(6), pp. 639-650.

Keator, D. B., Grethe, J. S., Marcus, D., Ozyurt, B., Gadde, S., Murphy, S., Pieper, S., Greve, P. D., Notestine, R., Bockholt, H. J., *et al.* (2008). A national human neuroimaging collaboratory enabled by the biomedical informatics research network (BIRN). *IEEE Transactions on Information Technology in Biomedicine*, 12(2), pp. 162-172.

Kern, L. M. & Kaushal, R. (2007). Health information technology and health information exchange in New York State: New initiatives in implementation and evaluation. *Journal of Biomedical Informatics*, 40(6 SUPPL.), pp. S17-S20.

Knaup, P., Bott, O., Kohl, C., Lovis, C. & Garde, S. (2007). Electronic patient records: moving from islands and bridges towards electronic health records for continuity of care. *Yearbook Of Medical Informatics*, pp. 34-46.

Kohane, I. S. (2000). Bioinformatics and Clinical Informatics: The Imperative to Collaborate. *J Am Med Inform Assoc*, 7(5), pp. 512-516.

Kohn, L. T., Corrigan, J. & Donaldson, M. S. (eds.) (2000). *To err is human: building a safer health system,* Washington, D.C.: National Academies Press.

Kripalani, S., LeFevre, F., Phillips, C. O., Williams, M. V., Basaviah, P. & Baker, D. W. (2007). Deficits in communication and information transfer between hospital-based and primary care physicians: Implications for patient safety and continuity of care. *Journal of the American Medical Association*, 297(8), pp. 831-841.

Kuhn, K. A., Wurst, S. H. R., Bott, O. J. & Giuse, D. A. (2006). Expanding the scope of health information systems. Challenges and developments. *Yearbook Of Medical Informatics*, pp. 43-52.

Kukafka, R., Ancker, J. S., Chan, C., Chelico, J., Khan, S., Mortoti, S., Natarajan, K., Presley, K. & Stephens, K. (2007). Redesigning electronic health record systems to support public health. *Journal of Biomedical Informatics*, 40(4), pp. 398-409.

Lagouarde, P., Blanc, B. & Vallespir, B. (2007). Interoperability in Healthcare: Standards for the French Project of a National Personal Medical Record. *In:* Doumeingts, G., Müller, J., Morel, G. & Vallespir, B. (eds.) *Enterprise Interoperability*. Springer London.

Larman, C. (2003). *Agile and Iterative Development: A Manager's Guide*, Addison-Wesley Professional.

Lenz, R., Beyer, M. & Kuhn, K. A. (2007). Semantic integration in healthcare networks. *International Journal of Medical Informatics*, 76(2-3), pp. 201-207.

Lewis, W. E., Dobbs, D. & Veerapillai, G. (2008). *Software Testing and Continuous Quality Improvement*, Taylor & Francis.

Leymann, F., Roller, D. & Schmidt, M. T. (2002). Web services and business process management. *IBM Systems Journal*, 41(2), pp. 198-211.

Littlejohns, P., Wyatt, J. C. & Garvican, L. (2003). Evaluating computerised health information systems: hard lessons still to be learnt. *BMJ: British Medical Journal*, 326(7394), p. 860.

Liu, X., Shi, Y., Xu, Y., Tian, Y. & Liu, F. (2010). Heterogeneous Database Integration of EPR System Based on OGSA-DAI. *In:* Zhang, W., Chen, Z., Douglas, C. & Tong, W. (eds.) *High Performance Computing and Applications*. Springer Berlin / Heidelberg.

LOINC (2012). *Logical Observation Identifiers Names and Codes* [Online]. Available: http://loinc.org/ [Accessed 2012-02-23].

Lopes, P. & Oliveira, J. L. (2011). COEUS: A Semantic Web application framework. Proceedings of the 4th International Workshop on Semantic Web Applications and Tools for the Life Sciences (SWAT4LS '11), 2011 London, UK. ACM, 66-73.

López, D. M. & Blobel, B. (2010). Architectural approaches for HL7-based health information systems implementation. *Methods of Information in Medicine*, 49(2), pp. 196-204.

Lopez, D. M. & Blobel, B. G. M. E. (2009). A development framework for semantically interoperable health information systems. *International Journal of Medical Informatics*, 78(2), pp. 83-103.

MAeHC (2012). *Massachusetts eHealth Collaborative* [Online]. Available: http://www.maehc.org/ [Accessed 2012-02-23].

Mäenpää, T., Asikainen, P., Gissler, M., Siponen, K., Maass, M., Saranto, K. & Suominen, T. (2011). Outcomes Assessment of the Regional Health Information Exchange. *Methods of Information in Medicine*, 50(4), pp. 308-318.

Mäenpää, T., Asikainen, P., Gissler, M., Siponen, K., Maass, M., Saranto, K. & Suominen, T. (2012). The utilization rate of the regional health information exchange: How it impacts on health care delivery outcomes. *Journal of Public Health Management and Practice*, 18(3), pp. 215-223.

Mäenpää, T., Suominen, T., Asikainen, P., Maass, M. & Rostila, I. (2009). The outcomes of regional healthcare information systems in health care: A review of the research literature. *International Journal of Medical Informatics*, 78(11), pp. 757-771.

Maffei, R., Burciago, D. & Dunn, K. (2009). Determining business models for financial sustainability in regional health information organizations (RHIOs): a review. *Population health management*, 12(5), pp. 273-278.

Markle-Foundation (2012). *Markle Connecting for Health : Markle Common Framework* [Online]. Available: http://www.markle.org/health/markle-common-framework [Accessed 2012-03-30].

Maro, J. C., Platt, R., Holmes, J. H., Strom, B. L., Hennessy, S., Lazarus, R. & Brown, J. S. (2009). Design of a National Distributed Health Data Network. *Ann Intern Med*, 15(5), pp. 341-344.

Martin-Sanchez, F., Iakovidis, I., Nørager, S., Maojo, V., De Groen, P., Van Der Lei, J., Jones, T., Abraham-Fuchs, K., Apweiler, R., Babic, A., *et al.* (2004). Synergy between medical informatics and bioinformatics: Facilitating genomic medicine for future health care. *Journal of Biomedical Informatics*, 37(1), pp. 30-42.

McCay, C., Kalra, D. & Worden, R. (2008). *Results of Investigating the Transformability Between HL7 V3, openEHR and EN/ISO 13606*, NHS Connecting for Health (Report NPFIT-FNT-TO-SCG-0025.02).

McConnell, S. (2004). *Code complete*, Microsoft Press.

McDonald, C. J., Overhage, J. M., Barnes, M., Schadow, G., Blevins, L., Dexter, P. R. & Mamlin, B. (2005). The Indiana network for patient care: a working local health information infrastructure. *Health Aff (Millwood)*, 24(5), pp. 1214-20.

McDonald, C. J., Overhage, J. M., Dexter, P., Takesue, B. & Suico, J. G. (1998). What is done, what is needed and what is realistic to expect from medical informatics standards. *International Journal of Medical Informatics*, 48(1-3), pp. 5-12.

Medcom (2012). *Medcom - det danske sundhedsdatanet* [Online]. Available: http://www.medcom.dk/ [Accessed 2012-02-23].

Microsoft (2009). *Connected Health Framework Architecture and Design Blueprint (2nd ed)* [Online]. Microsoft. Available: http://www.microsoft.com/health/ww/ict/Pages/Connected-Health-Framework.aspx [Accessed 2012-01-20].

Mildenberger, P., Eichelberg, M. & Martin, E. (2002). Introduction to the DICOM standard. *European Radiology*, 12(4), pp. 920-927.

Miller, R. H. & Miller, B. S. (2007). The Santa Barbara County Care Data Exchange: What happened? *Health Affairs*, 26(5), pp. 568-580.

Mitchell, J. A., McCray, A. T. & Bodenreider, O. (2003). From Phenotype to Genotype: Issues in Navigating the Available Information Resources. *Methods of Information in Medicine*, 42(5), pp. 557-563.

Montagnat, J., Breton, V. & Magnin, I. E. (2005). Partitioning medical image databases for content-based queries on a grid. *Methods of Information in Medicine*, 44(2), pp. 154-160.

Montagnat, J., Frohner, Á., Jouvenot, D., Pera, C., Kunszt, P., Koblitz, B., Santos, N., Loomis, C., Texier, R., Lingrand, D., *et al.* (2008a). A Secure Grid Medical Data Manager Interfaced to the gLite Middleware. *Journal of Grid Computing*, 6(1), pp. 45-59.

Montagnat, J., Gaignard, A., Lingrand, D., Balderrama, J. R., Collet, P. & Lahire, P. (2008b). NeuroLOG: a community-driven middleware design. *In:* Solomonides, T., Silverstein, J. C., Saltz, J., *et al.* (eds.) *Global Healthgrid: E-Science Meets Biomedical Informatics*. Amsterdam: IOS Press.

Monteagudo, J. L. & Moreno, O. (2007). *eHealth strategy and implementation activities in Portugal*, eHealth ERA Project (Towards the Establishment of a European Research Area). Available: http://www.ehealth-era.org/database/documents/ERA_Reports/Portugal_eHealth-ERA_country_report_final_01-06-2007.pdf.

Mykkänen, J., Riekkinen, A., Sormunen, M., Karhunen, H. & Laitinen, P. (2007). Designing web services in health information systems: From process to application level. *International Journal of Medical Informatics*, 76(2-3), pp. 89-95.

NAfHIT (2008). *Report to the Office of the National Coordinator for Health Information Technology on defining key health information technology terms*, The National Alliance for Health Information Technology.

Nygren, E. & Henriksson, P. (1992). Reading the medical record. I. Analysis of physicians' ways of reading the medical record. *Computer Methods and Programs in Biomedicine*, 39(1-2), pp. 1-12.

OASIS (2012). *OASIS Standards* [Online]. Available: http://www.oasis-open.org/standards [Accessed 2012-02-23].

Oates, J. & Jensen, H. B. (eds.) (2000). *Building Regional Health Care Networks in Europe*: IOS Press.

OECD-HD (2011). *OECD Health Data 2011*. Available: http://www.oecd.org/health/healthdata [Accessed 2012-02-23].

OECD-HPS (2010). *Improving Health Sector Efficiency : The Role of Information and Communication Technologies*, OECD Health Policy Studies. Available: http://dx.doi.org/10.1787/9789264084612-en [Accessed 2012-02-23].

OGF (2012). Open Grid Services Architecture (OGSA) Primer. Available: http://www.gridforum.org/OGSA_Primer/ [Accessed 2012-02-23].

Oh, H., Rizo, C., Enkin, M. & Jadad, A. (2005). What is eHealth (3): A systematic review of published definitions. *Journal of Medical Internet Research*, 7(1).

Olabarriaga, S. D. & Montagnat, J. (2010). Special section: Medical imaging on grids. *Future Generation Computer Systems*, 26(4), pp. 678-680.

Oliveira, I. C., Alves, L., Dias, E., Pacheco, D., Lima, S., Barros, J., Monteiro, M. P., Silva, J. A., Fernandes, J. M., Cunha, J. P. S., *et al.* (2010). Automated endoscopic capsule analysis using a Grid computing environment. *In:* Proenca, A., Pina, A., Tobio, J. G. & Ribeiro, L., eds. Ibergrid: 4th Iberian Grid Infrastructure Conference Proceedings, 2010. Netbiblo, pp. 319-330.

Oliveira, I. C., Belo, O. & Cunha, J. P. S. (2000). Agents working on the integration of heterogeneous information sources in distributed healthcare environments. *In:* Monard, M. C. & Sichman, J. S. (eds.) *International Joint Conference of the 7th Ibero-American Conference on AI/15th Brazilian Symposium on AI.* Sao Paulo, Brazil: Springer-Verlag Berlin: 136-145.

Oliveira, I. C., Cunha, J. P. S., Pacheco, D., Fernandes, J. M., Pedrosa, M., Alves, L. & Pereira, A. S. (2009). A neuroscience Grid-enabled portal for the Portuguese Brain Imaging Network. *In:* Frisoni, G., Barattieri, C. & Manset, D. (eds.) *MICCAI-Grid Workshop.* London, UK.

Oliveira, I. C., Dias, E., Alves, L., Barros, J., Silva, J. A., Monteiro, M. P., Pereira, A. S., Fernandes, J. M., Campilho, A. & Cunha, J. P. S. (2011). Extending a desktop endoscopic capsule video analysis tool used by doctors with advanced computing resources. *In:* Palencia, I. C., Lucas, J. M., Gomes, J. & Barreira, G., eds. 5th Iberian Grid Infrastructure Conference (IberGrid 2011), 2011 Santander (Spain). NetBiblio, pp. 156-168.

Oliveira, I. C., Fernandes, J. M., Alves, L., Pereira, A. S. & Cunha, J. P. S. (2008). GERES-med: An Architecture for Grid-Enabled scientific REpositorieS for medical applications. *In:* Silva, F., Barreira, G. & Ribeiro, L., eds. Ibergrid: 2nd Iberian Grid Infrastructure Conference Proceedings, 2008. Netbiblo, 163-173.

Oliveira, I. C., Oliveira, J. L., Sanchez, J. P., López-Alonso, V., Martin-Sanchez, F., Maojo, V. & Sousa Pereira, A. (2005). Grid requirements for the integration of biomedical information resources for health applications. *Methods of Information in Medicine*, 44(2), pp. 161-167.

OMG (2001). *Clinical Observation Access Service (COAS)*, OMG. Available: http://www.omg.org/spec/COAS/.

Ortega Egea, J. M., González, M. V. R. & Menéndez, M. R. (2010). eHealth usage patterns of European general practitioners: A five-year (2002-2007) comparative study. *International Journal of Medical Informatics*, 79(8), pp. 539-553.

Oster, S., Langella, S., Hastings, S., Ervin, D., Madduri, R., Phillips, J., Kurc, T., Siebenlist, F., Covitz, P., Shanbhag, K., *et al.* (2008). caGrid 1.0: An Enterprise Grid Infrastructure for Biomedical Research. *Journal of the American Medical Informatics Association*, 15(2), pp. 138-149.

Pagliari, C., Detmer, D. & Singleton, P. (2007). Potential of electronic personal health records. *British Medical Journal*, 335(7615), pp. 330-333.

Painter, J., Kirnak, A. & Moehrke, J. (2012). *A Service-Oriented Architecture (SOA) View of IHE Profiles*, Integrating the Healthcare Enterprise (IHE IT Infrastructure White Paper). Available: http://www.ihe.net/Technical_Framework/upload/IHE_ITI_TF_WhitePaper_A-Service-Oriented-Architecture_SOA_2009-09-28.pdf.

Papazoglou, M. P. & van den Heuvel, W. J. (2007). Service oriented architectures: approaches, technologies and research issues. *Vldb Journal*, 16(3), pp. 389-415.

Parent, C. & Spaccapietra, S. (1998). Issues and approaches of database integration. *Communications of the ACM*, 41(5).

Paul, D. (2007). *Continuous Integration : Improving Software Quality and Reducing Risk*, Pearson Education.

Pereira, A. S., Maojo, V., Martin-Sanchez, F., Babic, A. & Goes, S. (2002). The Infogenmed Project. *ICBME 2002.* Singapore.

Philips-EIU (2007). *The future of healthcare : Putting the patient at the centre of care*, Philips briefing paper written in conjunction with the Economist Intelligence Unit. Available: http://www.getinsidehealth.com/PageFiles/402/The%20future%20of%20healthcare.pdf [Accessed 2012-01-20].

Plaisant, C., Mushlin, R., Snyder, A., Li, J., Heller, D. & Shneiderman, B. (1998). LifeLines: Using visualization to enhance navigation and analysis of patient records. *Journal of the American Medical Informatics Association*, pp. 76-80.

Poissant, L., Pereira, J., Tamblyn, R. & Kawasumi, Y. (2005). The impact of electronic health records on time efficiency of physicians and nurses: A systematic review. *Journal of the American Medical Informatics Association*, 12(5), pp. 505-516.

Porter, M. E. & Teisberg, E. O. (2006). *Redefining Health Care: Creating Value-Based Competition on Results*, Harvard Business Press.

Power, D. J., Politou, E. A., Slaymaker, M. A. & Simpson, A. C. (2005). Towards secure Grid-enabled healthcare. *Software - Practice and Experience*, 35(9), pp. 857-871.

Progress-Consulting & Living-Prospects (2011). *Dynamic health systems and new technologies: eHealth solutions at regional and local level*, European Union - Comittee of the Regions (Study report).

PWC (2009). *Transforming healthcare through secondary use of health data*, PricewaterhouseCoopers LLP - Health Industries (white paper). Available: http://www.pwc.com/us/en/healthcare/publications/secondary-health-data.jhtml [Accessed 2012-01-20].

Qureshi, W. A. (2004). Current and future applications of the capsule camera. *Nature Reviews Drug Discovery*, 3(5), pp. 447-450.

RCGP-HIG (2009). *Informing shared clinical care: Final (reference) report of the Shared Record Professional Guidance*, Royal College of General Practitioners - Health Informatics Group (guidelines report). Available: http://www.rcgp.org.uk/health_informatics_group/srpg.aspx [Accessed 2012-02-23].

Remen, V. M. & Grimsmo, A. (2011). Closing information gaps with shared electronic patient summaries––How much will it matter? *International Journal of Medical Informatics*, 80(11), pp. 775-781.

Rosenthal, A., Mork, P., Li, M. H., Stanford, J., Koester, D. & Reynolds, P. (2010). Cloud computing: A new business paradigm for biomedical information sharing. *Journal of Biomedical Informatics*, 43(2), pp. 342-353.

RSE-WG (2009a). *RSE - R2A: Orientações para Especificação Funcional e Técnica do Sistema de RSE [Guidelines for the Functional and Technical Specification of the RSE System]*, Grupo de Trabalho para o Registo de Saúde Electrónico (RSE). Available: http://www.portaldasaude.pt/portal/conteudos/a+saude+em+portugal/politica+da+saude/discussao/arquivo/registo+electronico.htm.

RSE-WG (2009b). RSE – Registo de Saúde Electrónico: R1 Documento de Estado da Arte [State of the Art review on Electronic Healt Records in Portugal].

RTS (2005). *Relatório de Especificação Detalhada de Procedimentos [Detailed Process Specification Report]*, Projeto Rede Telemática da Saúde.

Safran, C., Bloomrosen, M., Hammond, W. E., Labkoff, S., Markel-Fox, S., Tang, P. C. & Detmer, D. E. (2007). Toward a National Framework for the Secondary Use of Health Data: An American Medical Informatics Association White Paper. *Journal of the American Medical Informatics Association*, 14(1), pp. 1-9.

Saltz, J., Oster, S., Hastings, S., Langella, S., Kurc, T., Sanchez, W., Kher, M., Manisundaram, A., Shanbhag, K. & Covitz, P. (2006). caGrid: Design and implementation of the core architecture of the cancer biomedical informatics grid. *Bioinformatics*, 22(15), pp. 1910-1916.

Santos, V., Oliveira, D., Oliveira, I. C. & Cunha, J. P. S. (2009). A monitoring toolkit for a distributed clinical data integration engine. *In:* Azevedo, L. & Londral, A. R., eds. 2nd International Conference on Health Informatics (HEALTHINF 2009), Jan 14-17 2009 Oporto, Portugal. 300-305.

Saranummi, N., Piggott, D., Katehakis, D. G., Tsiknakis, M. & Bernstein, K. (eds.) (2005). *Regional Health Economies and ICT Services: The PICNIC Experience*: IOS Press.

Schabetsberger, T., Ammenwerth, E., Andreatta, S., Gratl, G., Haux, R., Lechleitner, G., Schindelwig, K., Stark, C., Vogl, R., Wilhelmy, I., *et al.* (2006). From a paper-based transmission of discharge summaries to electronic communication in health care regions. *International Journal of Medical Informatics*, 75(3-4), pp. 209-215.

Schabetsberger, T., Wozak, F., Katt, B., Mair, R., Hirsch, B. & Hörbst, A. (2010). Implementation of a Secure and Interoperable Generic e-Health Infrastructure for Shared Electronic Health Records based on

IHE Integration Profiles. *In:* Safran, C., Reti, S. & Marin, H. F., eds. MEDINFO 2010 - Proceedings of the 13th World Congress on Medical Informatics, 2010.

Schwaber, K. & Beedle, M. (2001). *Agile Software Development with Scrum*, Prentice Hall.

Serrano, J. A. V. & Kirchhoff, U. (2000). A Novel Methodology for World-best Teamwork in Hospitals. *In:* Stanford-Smith, B. & Kidd, P. T. (eds.) *E-business: Key issues, Applications and Technologies.* IOS Press.

Shaller, D. (2007). *Patient-centered care: what does it take?*, Report prepared for The Commonwealth Fund.

Sheth, A. P. & Larson, J. A. (1990). Federated database systems for managing distributed, heterogeneous, and autonomous databases. *ACM Computing Surveys*, 22(3), pp. 183-236.

SIENA (2012). *The SIENA Roadmap on Distributed Computing Infrastructure for e-Science and Beyond in Europe* Standards and Interoperability for eInfrastructure implemeNtation initiative (SIENA) -- European Comission. Available: http://www.sienainitiative.eu/ [Accessed 2012-02-23].

Simões, J. (ed.) (2010). *30 Anos do Serviço Nacional de Saúde - Um Percurso Comentado*: Almedina.

Simon, S. R., Evans, J. S., Benjamin, A., Delano, D. & Bates, D. W. (2009). Patients' Attitudes Toward Electronic Health Information Exchange: Qualitative Study. *J Med Internet Res*, 11(3).

SNOMED (2012). *SNOMED Clinical Terms* [Online]. Available: http://www.ihtsdo.org/snomed-ct/ [Accessed 2012-02-23].

Stead, W. W. & Lin, H. S. (eds.) (2009). *Computational Technology for Effective Health Care : Immediate Steps and Strategic Directions*: Committee on Engaging the Computer Science Research Community in Health Care Informatics, The National Academy Press.

Stroetmann, K. A., Jones, T., Dobrev, A. & Stroetmann, V. N. (2006). *eHealth is Worth it: The economic benefits of implemented eHealth solutions at ten European sites.* Available: http://www.ehealth-impact.org/download/documents/ehealthimpactsept2006.pdf [Accessed 2012-01-20].

Stroetmann, V. N., Kalra, D., Lewalle, P., Rector, A., Rodrigues, J. M., Stroetmann, K. A., Surjan, G., Ustun, B., Virtanen, M. & Zanstra, P. E. (2009). *Semantic Interoperability for Better Health and Safer Healthcare*, European Comission. Available: http://ec.europa.eu/information_society/activities/health/docs/publications/2009/2009semantic-health-report.pdf.

Tang, P. C., Ash, J. S., Bates, D. W., Overhage, J. M. & Sands, D. Z. (2006a). Personal health records: Definitions, benefits, and strategies for overcoming barriers to adoption. *Journal of the American Medical Informatics Association*, 13(2), pp. 121-126.

Tang, P. C. & McDonald, C. J. (2006b). Electronic Health Record Systems. *In:* Shortliffe, E. H. & Cimino, J. J. (eds.) *Biomedical Informatics.* Springer New York.

The 2020 Science Group (2005). *Towards 2020 Science*, Microsoft Research. Available: http://research.microsoft.com/en-us/um/cambridge/projects/towards2020science/ [Accessed 2012-02-23].

Toga, A. W. (2002). Neuroimage databases: the good, the bad and the ugly. *Nature Reviews. Neuroscience*, 3(4), pp. 302-309.

Tsiknakis, M., Brochhausen, M., Nabrzyski, J., Pucacki, J., Sfakianakis, S. G., Potamias, G., Desmedt, C. & Kafetzopoulos, D. (2008). A semantic grid infrastructure enabling integrated access and analysis of multilevel biomedical data in support of postgenomic clinical trials on cancer. *IEEE Transactions on Information Technology in Biomedicine*, 12(2), pp. 205-217.

Tsiknakis, M., Katehakis, D. G. & Orphanoudakis, S. C. (2002). An open, component-based information infrastructure for integrated health information networks. *International Journal of Medical Informatics*, 68(1-3), pp. 3-26.

UMIC (2010). *A Sociedade da Informação em Portugal 2010*, UMIC – Agência para a Sociedade do Conhecimento. Available: http://www.umic.pt/images/stories/publicacoes4/SIP2010_completo.pdf [Accessed 2012-02-23].

US-GAO (1993). *Automated Medical Records: Leadership needed to expedite standards development (report to the Chairman, Committee on Governmental Affairs, US Senate)* Comitee on Governamental Affairs -- United States Senate (GAO/IMTEC-93-17).

Uslu, A. M. & Stausberg, J. (2008). Value of the electronic patient record: An analysis of the literature. *Journal of Biomedical Informatics*, 41(4), pp. 675-682.

van de Velde, R. (1992). *Hospital information systems: the next generation*, Springer-Verlag.

van der Linden, H., Kalra, D., Hasman, A. & Talmon, J. (2009). Inter-organizational future proof EHR systems: A review of the security and privacy related issues. *International Journal of Medical Informatics*, 78(3), pp. 141-160.

Velez Lapão, L. (2007). Survey on the status of the hospital information systems in Portugal. *Methods of Information in Medicine*, 46(4), pp. 493-499.

Venugopal, S., Buyya, R. & Ramamohanarao, K. (2006). A taxonomy of data grids for distributed data sharing, management, and processing. *ACM Computing Surveys*, 38(1), pp. 123-175.

W3C *SOAP Specifications* [Online]. Available: http://www.w3.org/TR/soap/ [Accessed 2012-02-23].

W3C *Web Services* [Online]. Available: http://www.w3.org/2002/ws/ [Accessed 2012-02-23].

W3C (2000). *Extensible Markup Language (XML) 1.0* [Online]. The World Wide Web Consortium (W3C). Available: http://www.w3.org/TR/REC-xml [Accessed 2001-12-15].

W3C (2012). *World Wide Web Consortium (W3C) Standards and Drafts* [Online]. Available: http://www.w3.org/TR/.

Walker, J., Pan, E., Johnston, D., Adler-Mllstein, J., Bates, D. W. & Middleton, B. (2005). Market watch - The value of health care information exchange and interoperability. *Health Affairs*, 24(2), pp. W510-W518.

Wang, T. D., Wongsuphasawat, K., Plaisant, C. & Shneiderman, B. (2010). Visual information seeking in multiple electronic health records: design recommendations and a process model. *Proceedings of the 1st ACM International Health Informatics Symposium.* Arlington, Virginia, USA: ACM: 46-55.

Warren, R., Solomonides, A. E., del Frate, C., Warsi, I., Ding, J., Odeh, M., McClatchey, R., Tromans, C., Brady, M., Highnam, R., *et al.* (2007). MammoGrid - a prototype distributed mammographic database for Europe. *Clinical Radiology*, 62(11), pp. 1044-1051.

Watch, T. E. e.-B. M. (2005). CASE STUDY: MEDCOM - THE DANISH HEALTHCARE DATA AND INFORMATION NETWORK 1. *e-Business Watch* [Online]. Available: http://ec.europa.eu/enterprise/archives/e-business-watch/studies/case_studies/documents/Case%20Studies%202004/CS_SR10_Health_1-MedCom.pdf.

WHO-ROE (2010). *Portugal Health System Performance Assessment*, World Health Organization Regional Office for Europe and Ministry of Health of Portugal Available: http://www.euro.who.int/en/what-we-do/health-topics/Health-systems/health-systems-governance/publications/2010/portugal-health-system-performance-assessment-2010 [Accessed 2012-02-23].

WHO (2003). *International Classification of Primary Care, Second edition (ICPC-2)* [Online]. World Health Orgnization. Available: http://www.who.int/classifications/icd/adaptations/icpc2/en/index.html.

WHO (2008). *Building Foundations for eHealth in Europe*, WHO Global Observatory for eHealth. Available: http://www.who.int/goe/BFeuroFull.pdf.

Winter, A. (2009). The future of medical informatics - Some perspectives of intra- and inter-institutional information systems. *Methods of Information in Medicine*, 48(1), pp. 62-65.

Wöhrer, A., Brezany, P. & Tjoa, A. M. (2005). Novel mediator architectures for Grid information systems. *Future Generation Computer Systems*, 21(1), pp. 107-114.

Wollersheim, D., Sari, A. & Rahayu, W. (2009). Archetype-based electronic health records: a literature review and evaluation of their applicability to health data interoperability and access. *Health Information Management Journal*, 38(2), pp. 7-17.

Wyatt, J. (1994). Clinical data systems, part 1: data and medical records. *The Lancet*, 344(8936), pp. 1543-1547.

Yang, T. H., Sun, Y. L. S. & Lai, F. P. (2011). A Scalable Healthcare Information System Based on a Service-oriented Architecture. *Journal of Medical Systems*, 35(3), pp. 391-407.

Zuquete, A., Gomes, H. & Cunha, J. P. S. (2008). Authentication of professionals in the RTS e-Health system. *In:* Azevedo, L. & Londral, A. R., eds. Healthinf 2008: Proceedings of the First International Conference on Health Informatics, Vol 1, 2008 Setubal. Insticc-Inst Syst Technologies Information Control & Communication, 72-80.